Technical Report 1089

Virtual Environments for Dismounted Soldier Training and Performance: Results, Recommendations, and Issues

Bruce W. Knerr, Donald R. Lampton, Michael J. Singer, Bob G. Witmer, and Stephen L. Goldberg
U.S. Army Research Institute

Kimberly J. Parsons and James Parsons
Institute for Simulation and Training

November 1998

United States Army Research Institute for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

A Directorate of the U.S. Total Army Personnel Command

EDGAR M. JOHNSON
Director

Technical Review by

Stuart Grant, Defence and Civil Institute of Environmental Medicine, Canada
David Ryan-Jones, Naval Air Warfare Center Training Systems Division

NOTICES

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) November 1998	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) October, 1991 - September, 1998
4. TITLE AND SUBTITLE Virtual Environments for Dismounted Soldier Training and Performance: Results, Recommendations, and Issues		5a. CONTRACT OR GRANT NUMBER N/A
		5b. PROGRAM ELEMENT NUMBER 2O262785A
6. AUTHOR(S) Bruce W. Knerr, Donald R. Lampton, Michael J. Singer, Bob G. Witmer, & Stephen L. Goldberg (U.S. Army Research Institute); Kimberly J. Parsons and James Parsons (Institute for Simulation and Training)		5c. PROJECT NUMBER A791
		5d. TASK NUMBER 2111
		5e. WORK UNIT NUMBER H01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: TAPC-ARI-IF 12350 Research Parkway Orlando, FL 32826-3276		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI
		11. MONITOR REPORT NUMBER Technical Report 1089

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*: The U.S. Army has made a considerable investment in the use of virtual environments (VE) to train combat forces, to evaluate new systems and operational concepts, and to rehearse specific missions. While these simulations have predominately focused on training and simulation for mounted soldiers, there is also a need to train infantry and other dismounted soldiers. Although VEs have the potential to immerse dismounted soldiers directly in simulations, there are few successful examples of the use of VE to provide effective training. The effective use of VE for training requires identification of the types of tasks for which VE training is most appropriate, the characteristics of VE systems that are required to provide effective training, and the training strategies that are most appropriate for use with VE. This report presents recommendations for the use of VE for dismounted soldier training and mission rehearsal, and identifies needed future research. They are based on the results of an ARI in-house research program, related programs in which ARI scientists have participated, and the work of other VE researchers. Recommendations include types of tasks for which training in VE is and is not appropriate, interface design recommendations, and ways to reduce side- and after-effects.

15. SUBJECT TERMS

Virtual Environments, Virtual Reality, Dismounted Infantry, Training, Presence, Simulator Sickness, Training Transfer

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	80	Dr. Bruce W. Knerr (407) 384-3987

Technical Report 1089

Virtual Environments for Dismounted Soldier Training and Performance: Results, Recommendations, and Issues

Bruce W. Knerr, Donald R. Lampton, Michael J. Singer, Bob G. Witmer, and Stephen L. Goldberg
U.S. Army Research Institute

Kimberly J. Parsons and James Parsons
Institute for Simulation and Training

Simulator Systems Research Unit
Stephen L. Goldberg, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

November 1998

Army Project Number
20262785A791

Education and Training Technology

FOREWORD

The U.S. Army has made a substantial commitment to the use of simulation for training, readiness, concept development, and test and evaluation. The current virtual simulation training system, the Close Combat Tactical Trainer, is designed to provide realistic simulation of platform-based warfighting, not to provide realistic training for dismounted soldiers. Virtual Environment (VE) technology, which uses position tracking and real-time update of visual, auditory, and other displays (e.g., tactile) has the potential to improve simulation-based training for dismounted soldiers. Since VE technologies are so new, there is little information to indicate where and how they are best used for dismounted soldier training.

This report presents recommendations for the use of VE for dismounted soldier training and mission rehearsal, and identifies needed future research. Recommendations include types of tasks for which training in VE is and is not appropriate, interface design recommendations, and ways to reduce side- and after-effects. The recommendations are based on the results of an ARI in-house research program, related programs in which ARI scientists have participated, and the work of other VE researchers.

The U.S. Army Research Institute, Simulator Systems Research Unit, conducts research with the goal of improving the effectiveness of simulators and simulations. The work described here is a part of ARI Research Task 2111, Virtual Environments for Combat Training and Mission Rehearsal.

Zita M. Simutis

ZITA M. SIMUTIS
Technical Director

ACKNOWLEDGEMENTS

Our thanks go to the present and past Consortium Research Fellows listed below. These University of Central Florida students ran experiments, collected and analyzed data, wrote papers, and contributed to the reports that described our program. This research would not have been possible without their efforts.

Robert Allen
John Bailey
James Bliss
Steven Cinq-Mars
Jennifer Ehrlich
John Gildea
Brian Kline
Eugenia Kolasinski
Daniel McDonald
Mara Rodriguez

VIRTUAL ENVIRONMENTS FOR DISMOUNTED SOLDIER TRAINING AND PERFORMANCE: RESULTS, RECOMMENDATIONS, AND ISSUES

EXECUTIVE SUMMARY

Research Requirement:

The U.S. Army has made a considerable investment in the use of virtual environments (VE) for training personnel who fight from within combat vehicles. The U.S. Army also needs a simulation capability to: train dismounted soldiers to operate with mounted units; train dismounted leaders, teams, and individuals to perform unit tasks; and plan and rehearse specific missions. The effective use of VE for training requires identification of the types of tasks for which VE training is most appropriate, the characteristics of VE systems that are required to provide effective training, and the training strategies that are most appropriate for use with VE.

Procedure:

In order to develop this knowledge, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI), Simulator Systems Research Unit, initiated a program of in-house experimentation to investigate the use of VE technology to train dismounted soldiers in 1992. Following an initial analysis of the task requirements for dismounted soldier training, and a review of the previous research in the use of VE for training, we conducted four experiments to investigate interface effects on the capabilities of participants to perform simple tasks in VE. We conducted two experiments that addressed the effectiveness of VE for teaching the configuration of and routes through large buildings, and the transfer of the knowledge acquired to the real world. These results led to a program of basic research into the problems of distance estimation in VE. We investigated the use of VE to represent exterior terrain, both for training land navigation skills (identifying landmarks and learning routes), and transfer to the actual terrain. Finally, we are investigating the use of VE for training team tasks. Overall, we have conducted 13 experiments involving over 500 human subjects. The results of our research program have been supplemented by the results of related programs in which we have participated, and a review related VE research.

Findings:

VE can be used effectively to train spatial knowledge. It is not currently suitable for training soldier tasks that involve precise or rapid motor activity. This includes training tasks that require accurate use of individual weapons. Systems that use Helmet-Mounted Displays (HMDs) are not recommended for tasks that require rapid head movements. Relatively inexpensive HMDs provide poor visual resolution, and are not recommended if successful task performance or training require fine visual discriminations, such as identifying targets at long distances. Distance estimation in VE is poor. Tasks requiring accurate estimation of distance require the

use of a display with a large field of view (FOV). Compensatory cues or some similar method for calibrating distance judgments in the VE may improve distance estimates.

The process of moving through the VE should require minimal attention on the part of the user. If the task requires use of the hands to perform other tasks simultaneously or nearly simultaneously with movement, then some form of foot-based locomotion control will be required. Otherwise, a well-practiced method that is not like the real world (such as a joystick) may be preferable to a novel method which is more similar to actual walking. Trainees should always receive training and practice in the use of the interface before they begin training on the target task or tasks.

Simulator sickness is a real but manageable problem. Its reduction should be a major consideration in the design and development of a VE-based training system, and in determining the environment, activities, and duration of the trainee's first use of the VE.

A number of unresolved research issues were also identified. These issues are also primarily concerned with improving VE interfaces, improving VE training effectiveness and transfer, and reducing simulator sickness symptoms.

Utilization of Findings:

These recommendations will be used to select the tasks and design the soldier-computer interfaces for prototype VE systems for dismounted infantry training. The research issues identified will be used to establish priorities for future research.

VIRTUAL ENVIRONMENTS FOR DISMOUNTED SOLDIER TRAINING AND PERFORMANCE: RESULTS, RECOMMENDATIONS, AND ISSUES

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

Virtual Environments for Dismounted Soldier Training and Performance: Results, Recommendations, and Issues

I. Introduction

Background

The U.S. Army has made a considerable investment in the use of virtual environments (VE) for training. Beginning in the early 1980's with the SIMNET program, and continuing with the Close Combat Tactical Trainer, the U.S. Army has made a commitment to use virtual simulations to train combat forces and to evaluate new systems and operational concepts. These simulations have predominantly focused on training personnel who fight from within combat vehicles (e.g., tanks and infantry fighting vehicles). More recently, the need to train infantry and other dismounted soldiers in these simulations has been recognized.

Members of small dismounted infantry (DI) units will face growing responsibilities and increasing challenges in combined arms combat and in contingency operations on the digitized battlefield of the Army After Next. These multiple challenges include new missions, changing doctrine, new and increasingly sophisticated equipment, and the need to fight with mounted units.

The increasing need to perform operations that require skills other than combat is illustrated by recent actions in Haiti, Somalia, and Bosnia. Non-combat missions involve more diverse types of operations and more complex and dynamic rules of engagement than do combat missions. When combined with dispersed units and a high level of media attention, this creates an environment in which the actions and decisions of small-unit leaders and individual soldiers can have far-reaching impacts. Such operations tend to make the job of small unit leaders (fire team, squad, and platoon leaders) more difficult by separating units, limiting visibility, and producing combat at close ranges. Field training for these diverse missions is limited by time, cost and safety factors.

Evolving doctrine, as illustrated by Force XXI and the Army After Next, together with digital communications systems and computerized weapons systems, will shift information and decision-making further down the chain of command. Soldiers will have to become more capable of either autonomous or small group action. Therefore, soldiers will also need to be more adaptive. "Leadership at the lowest level" will become more important. Increased flexibility and adaptability will be required at all levels of leadership. TRADOC's Army After Next, which focuses on warfare between the years 2010 and 2025, envisions squad leaders making decisions that could have strategic national importance (Naylor, 1996).

A variety of new equipment is being proposed, developed, and evaluated to improve the effectiveness and survivability of these dismounted units. This includes the soldier computer, global positioning system, squad radio, digital information display, video capture capability,

night vision equipment, and equipment for enhanced Military Operations in Urban Terrain (MOUT) capability (Soldier Systems Command, 1998). More sophisticated equipment will permit lower echelons to have more information about the tactical situation, command of more firepower, and greater physical separation from their parent unit. It should also lead to increased independence of action and decision-making responsibility at low echelons. However, these capabilities require new tactics for their use, as well as new skills on the part of the soldiers who must use them.

As a result of these factors, the U.S. Army needs improved capabilities for dismounted soldier simulation to support the following activities:

- Training dismounted leaders, teams, and individuals to perform unit tasks;
- Training dismounted soldiers to operate with mounted units;
- Planning and rehearsal of specific missions;
- Conducting human-in-the-loop experiments to develop requirements for new dismounted soldier concepts and technologies; and
- Test and evaluation of dismounted soldier concepts and equipment prototypes.

Although technologies such as helmet-mounted visual displays, head trackers, 3-D sound systems, haptic devices, and powerful graphics image generators have the potential to immerse dismounted soldiers directly in virtual training environments, their capability to provide effective training has yet to be ascertained. The effective use of VE for training requires more than just VE hardware and software. It also requires a body of knowledge that identifies: (a) the types of tasks for which VE training is most appropriate; (b) the characteristics of VE systems that are required to provide effective training; and (c) the training strategies that are most appropriate for use with VE. In order to develop this body of knowledge, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI), Simulator Systems Research Unit, initiated a program of experimentation to investigate the use of VE technology to train dismounted soldiers in 1992. We have now reached the point in our program where it is possible to synthesize what we and other researchers have found, draw initial conclusions, and make recommendations for the use of VE for U.S. Army training.

Report Objective

The objective of this report is to summarize the current state of knowledge of the use of VE for dismounted solder training. The core of this information comes from our in-house research program. Also included will be information which we have obtained from our participation in other, related programs. These include the assessment of simulator sickness in the M1 Tank Driver Trainer (M1TDT), the Individual Combatant Simulation System (ICSS), and the Dismounted Warrior Network (DWN). Other VE research has also been included. This research is synthesized and integrated to produce initial recommendations for the use of VE for dismounted soldier training and mission rehearsal, and to identify future research needs.

Overview and History of the Research Project

Gorman (1990) was an early proponent of the use of VE for DI training. His report called for the development of an individual portal (or I-Port) into VE. Partly as a result of his efforts, a conference was held in Snowbird, Utah in the fall of 1990 to discuss individual soldier systems and the role that an I-Port would play in their development (Goldberg and Knerr, 1997). While consensus was achieved on the need for an I-Port, Operation Desert Storm prevented the initiation of a cooperative effort.

The conference did provide the impetus for the initiation of an ARI program, however. The effort began with an initial examination of the feasibility of using VE technology for dismounted soldier training and the identification of difficult technical problems and research issues (Levison & Pew, 1993). This was followed up shortly thereafter with a more detailed examination of DI unit tasks and expected VE capabilities (Jacobs et al., 1994).

With these reports as a basis, ARI planned an in-house research program to investigate critical behavioral science research issues involved in dismounted soldier simulation. The research program, as it was initially planned, had a number of levels that could be thought of as taking the shape of a pyramid (see Figure 1). The progress of the program since its initial inception has stayed remarkably close to its original plan.

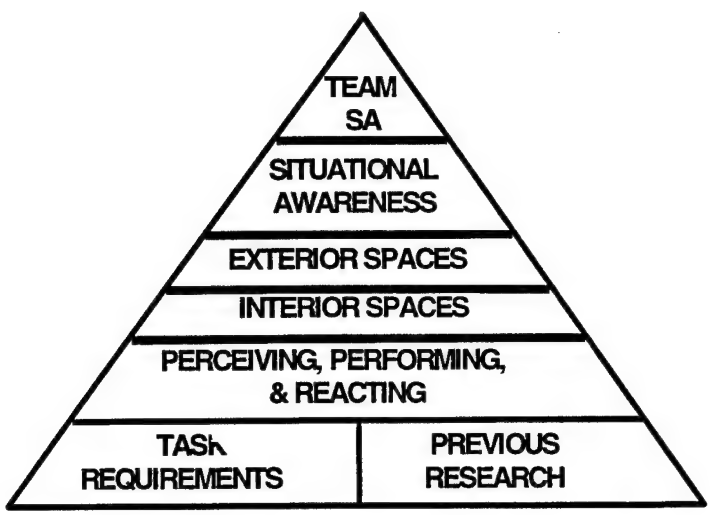

Figure 1. The virtual environment research pyramid.

Following an initial analysis of the task requirements for dismounted soldier training, and a review of the previous research in the use of VE for training (a very sparse area when we began), we conducted four experiments to investigate interface effects on the capabilities of participants to perform simple tasks in VE. In order to conduct these experiments we developed the Virtual Environment Performance Assessment Battery (VEPAB). Variables investigated included the type of control device, amount of practice on the tasks, stereoscopic vs. monoscopic HMDs, and type of display device (monitor, boom, HMD). At the next level, we conducted two

experiments that addressed the effectiveness of VE for teaching the configuration of and routes through large buildings, and the transfer of the knowledge acquired to the real world. Taken together, these results led to our proposing a program of basic research into the investigation of distance estimation in VE. At the third level, we investigated the use of VE to represent exterior terrain, both for training land navigation skills (identifying landmarks and learning routes), and performing threat assessments. Research at the top of the pyramid (we have combined the top two levels) is investigating the use of VE for training team tasks. Table 1 provides a summary of each of the experiments completed to date.

The core program was supplemented by our participation in a number of activities with other organizations to advance the state of knowledge. While not always directly related to dismounted soldier simulation, they nevertheless contributed to one or more of our research objectives.

Individual Combatant Simulation System (ICSS)

The ICSS concept was derived from the originally independent research and development efforts of the U.S. Air Force Armstrong Laboratory Human Resources Directorate, the U.S. Army Research Laboratory Human Research and Engineering Directorate, the U.S. Marine Corps Systems Command, the U.S. Naval Air Warfare Center, Training Systems Division, and ARI. In early 1994, with the encouragement of the Defense Modeling and Simulation Office, these organizations entered into a cooperative agreement for joint conduct of a program to enhance the ability of the Department of Defense to train individual combatant and leader skills, conduct virtual prototyping of developmental items, and provide the ability to conduct development and analysis of system utility, maintainability and human centered design. ARI participated in tasks involving human figure models and networking, voice and gesture control of semi-automated forces, and assessment planning. The program was terminated after one year for lack of funding.

M1 Tank Drive Trainer

At the request of TRADOC, the ARI Simulation Systems Research Unit and Armored Forces Research Units conducted a data collection and analysis effort to determine the incidence and severity of side effects resulting from use of the M1TDT, a motion base simulator. Recommendations were made for reducing their frequency and severity.

Dismounted Warrior Network (DWN)

DWN is an U.S. Army Simulation Training and Instrumentation Command (STRICOM) program to provide a reliable, low-cost, easy-to-use capability to insert dismounted soldiers into VE. During 1997, a series of engineering and user experiments were conducted to explore the utility of a DWN system as a research and analysis tool and to investigate different interfaces for inserting dismounted soldiers into virtual simulations. A joint government-contractor team selected Virtual Individual Combatant Simulators (VICS) based on three criteria: a desire to have a diverse mixture of characteristics to examine; a cost/benefit assessment of system characteristics; and expected system availability. Following VIC selection, performance and

Table 1

Summary of Experiments Conducted

No.	Short Title	Task	Major Independent Variables	N	Primary Reference
1.1	Virtual Environment Performance Assessment Battery (VEPAB)	Variety of visual, locomotion, & manipulation tasks	Control Device: joystick vs. spaceball	24	Lampton, Knerr, Goldberg, Bliss, Moshell, & Blau (1995)
1.2	Visual Display Type	7 VEPAB tasks: Distance Estimation, Search, Turns, Doorways, Bins, Choice Reaction Time, Tracking	Display device: BOOM, HMD, or Monitor	48	Lampton, Gildea, McDonald, & Kolasinski (1996)
1.3	VEPAB Extended Practice	5 VEPAB Tasks: Turns, Figure 8, Windows, Bins, Tracking	Practice session	6	Knerr, Goldberg, Lampton, Bliss, Moshell, & Blau (1994)
1.4	Stereoscopic vs. Monoscopic Displays	4 VEPAB tasks: Doorways, Bins, and Fixed & Moving Tracking	Stereoscopic vs. monoscopic display; Head coupling vs. no head coupling	48	Singer, Ehrlich, Cinq-Mars, & Papin (1995)
2.1	Spatial Learning & Transfer	Building Route learning and transfer; Building Survey Learning	Rehearsal medium (VE, real world, symbolic); Map vs. no map	60	Witmer et al. (1995), Witmer et al. (1996)
2.2	Spatial Learning Strategies	Building Route Learning and Survey Learning	VE learning instructions (Restricted vs. Exploratory); Head-tracked view vs. joystick view)	64	Bailey & Witmer (1994)
3.1	Terrain Appreciation	Indicating direction and distance to landmarks w/ feedback during training – no feedback during testing	Hi-level VE interface, Low-level VE interface, or Map study; Terrain databases (2)	48	Singer, Allen, McDonald, & Gildea (1997)
3.2	Terrain Transfer	Indicating direction and distance to landmarks w/ feedback during training – no feedback during testing	VE practice vs. map study in transfer to Terrain (Ft. Benning)	34	Singer, Allen McDonald, & Fober (1998)
4.1a	Distal Distance Estimates	Estimate perceived distance for far targets	Texture, Pattern, Relative Size	24	Witmer & Kline (1998)
4.1b	Proximal Distance Estimates	Estimate perceived distance for near targets	Field of View, Emergent Texture, Resolution	28	Kline & Witmer (1996)
4.2	Moving Distance Estimates	Estimate traversed distance and travel time	Movement Type, Movement Speed, Edge Rate, Compensatory Cues	72	Witmer & Kline (1998)
4.4	Non-visually Guided Locomotion (NVGL)	Walk the perceived distance to a target without vision	Environment Type, Environment Order, Performance Feedback	24	Witmer & Sadowski (in press)
5.1	Body Model	Search task during movement through offices	Presence or absence of body model	32	Singer, Ehrlich, & Allen (1998)

interoperability issues were identified and resolved. Finally, the selected VIC simulations, a DI Semi-automated Forces (SAF) station, an Exercise Support Station, and an After Action Review Station were tied into a distributed interactive simulation (DIS) network using DIS protocols. The ARI Simulator Systems Research Unit participated in the overall design of the experiments and conduct of the engineering experiments. The ARI Infantry Forces Research Unit participated in the conduct of the user experiments.

Organization of this Report

Following this introduction, Section II will examine the effects of the VE interface on human performance, with the emphasis on distance estimation, but considering other perceptual issues and motor task performance as well. Section III will discuss training in VE, to include skill acquisition, simulation fidelity and training effectiveness, transfer of skills acquired in VE to the real world, and instructional strategies on VE. Section IV looks at the concept of "presence" or sense of immersion in VE, and its impact on performance and learning. Section V is concerned with the problem of simulator sickness: causes, frequency and severity, and guidelines for prevention. Section VI discusses hardware and software issues in the use of VE. Section VII provides a summary of the recommendations for improving the effectiveness of the use of VE. Section VIII summarizes research issues that have been identified as requiring further research.

II. The Effects of Virtual Environment Interfaces On Task Performance

It is well known that the man-machine interface can affect human performance, but because VE interface technology is new, little is known about how it effects human performance. This section discusses the effects of VE interfaces on perceptual and psychomotor task performance, with emphasis on distance estimation tasks. Experiments related to distance judgments that were conducted in our laboratory are summarized in Table 2. The purpose of some of these experiments was to investigate the effects of manipulating critical variables and distance cues on the accuracy of distance judgments, while others looked at distance judgments as one component of spatial learning. The experiments differed not only in terms of their goals and objectives, but also in terms of the methods employed in obtaining the distance judgments. In addition to these experiments, we will briefly discuss the results from the DWN Experiments (Lockheed Martin, 1997; Pleban, Dyer, Salter, and Brown, 1997).

Distance Estimation

Methods Used

We employed several different techniques to evaluate the effects of our interface devices on the accuracy of VE distance estimates. Because each technique has its own strengths and weaknesses, the accuracy of the estimates obtained using the various techniques can vary dramatically. However, participants typically underestimate distance to a target in VEs, regardless of which technique is used.

Virtual Environment Performance Assessment Battery (VEPAB). Lampton et al. (1994) included a distance estimation task in ARI's VEPAB. The task was performed in a virtual office corridor with a checkerboard floor pattern and narrow vertical stripes on the walls every five feet. Participants estimated the distance to a stationary virtual soldier of known height positioned 40 feet from them. After reporting their estimate they were told the correct distance to the soldier. Then, as the soldier moved toward them at a constant speed, they reported when they perceived the soldier to be at distances of 30, 20, 10, 5, and 2.5 feet. The measure is the actual distance to the soldier when they report that the soldier to be 30, 20, 10, 5, and 2.5 feet away. Hence they are matching their distance perceptions to the designated egocentric distances. Although this method of obtaining distance estimates is easy to administer and can detect differences in estimation accuracy, it restricts the range of estimates that can be obtained, inflating the apparent accuracy of these estimates.

Projective Convergence. The projective convergence technique uses a participant's estimates of distance and bearing to a target not currently in the line-of-sight to determine the participant's perceived location of the target. The participants either draw lines of different lengths to represent the distance to targets or verbally report their distance judgments in standard or metric units. We used this technique to investigate spatial knowledge acquisition in learning building interiors (Witmer et al, 1995; Witmer, Bailey, Knerr, & Parsons, 1996; Bailey & Witmer, 1995) and in outside terrains (Singer, Allen, McDonald, & Gildea, 1997; Singer,

Table 2

Summary of Experiments Investigating Distance Judgments

Experiment No.	Authors	No. of Ss	Manipulated Variables/Distance Cues	Distance (feet)	Method Used to Obtain Distance Judgments
1.1	Lampton et al. (1994)	24 (Grp 1) 36 (Grp 2)	Type of Environment - VE Type of Environment -RW	2.5 to 40 2.5 to 40	Matching Egocentric Distances Matching Egocentric Distances
1.2	Lampton et al. (1995)	48 (Grp 1) 36 (Grp 2)	Type of Environment - VE, Display Device Type - VE Type of Environment - RW	2.5 to 40 2.5 to 40	Matching Egocentric Distances Matching Egocentric Distances
1.3	Lampton et al. (1996)	48 (Grp 1) 36 (Grp 2)	Type of Environment - VE, Display Device Type - VE Type of Environment - RW	2.5 to 40 2.5 to 40	Matching Egocentric Distances Matching Egocentric Distances
1.4	Singer et al. (1995)	48	Stereo vs. Monoscopic View, Head Tracking vs. Fixed View, Trials	2.5 to 40	Matching Egocentric Distances
2.1	Witmer et al. (1995), Witmer et al. (1996)	60	Rehearsal medium (VE, real world, symbolic), Map vs. no map	24 to 293	Projective Convergence (no feedback)
2.2	Bailey & Witmer (1994)	64	VE learning instructions (Restricted vs. Exploratory), (head-tracked views joystick view)	21 to 184	Projective Convergence (no feedback)
3.1	Singer et al. (1997)	48	Hi-VE (treadmill, Head-tracked view)versus Lo-VE (Joystick, Fixed View) versus Map Study	164 to 4920	Projective Convergence (feedback during training)
3.2	Singer et al. (1998)	34	Medium VE (Joystick, Head-tracked view) vs. Map Study, Test Site, Terrain	164 to 4920	Projective Convergence (feedback during training)
3.2	Allen & Singer (1997)	32	Hi-VE (treadmill, Head-tracked view) vs. Lo-VE (Joystick, Fixed View) vs. Map Study, Near (<1968 ft) versus Far (>1968 ft)	164 to 4920	Verbal Estimates (feedback during training)
4.1	Kline & Witmer (1996)	24	Field of View, Emergent Texture, Resolution	1 to 12	Magnitude Estimates
4.2	Witmer & Kline (1998)	28 (Exp 1) 72 (Exp 2)	Texture, Pattern, Relative Size Movement Type, Movement Speed, Edge Rate, Compensatory Cues	10 to 110 10 to 280	Magnitude Estimates Magnitude Estimates
4.4	Witmer & Sadowski (in press)	24	Environment Type, Environment Order, Performance Feedback	15 to 105	Non-visually Guide Locomotion (NVGL)

Allen, McDonald, & Fober, in preparation). Errors in estimated distance using this method may either be due to poor estimation of distance or lack of knowledge regarding the designated target location. Hence it is not a pure distance estimation measure.

Magnitude Estimation. Witmer & Kline (1998) and Kline & Witmer (1996) used a magnitude estimation technique for measuring distance perception. The task was performed in a virtual office corridor with varying floor and wall patterns and textures. Participants first estimated the distance to a standard stimulus (e.g., a cylinder at 100 feet). They received no feedback regarding the accuracy of their distance estimates to the standard stimulus, but were told that all subsequent estimates should be made relative to that standard. Actual distances varied from 1 to 12 feet in one experiment (Kline & Witmer, 1996), from 10 to 110 feet in another, and from 10 to 280 feet in a third (Witmer & Kline, 1998). The basic measure was the reported target distance in feet or meters. The ability of participants to accurately report distances in feet or meters varies dramatically among participants, and may be independent of their perception of target distance. These individual differences may inflate the amount of error observed in estimating target distance. The amount of error in these estimates was calculated as the difference between the estimated and true distance divided by the true distance. This error measurement is called relative error because it is the amount of error relative to the true target distance.

Non-Visually Guided Locomotion. Witmer & Sadowski (in press) used non-visually guided locomotion (NVGL) to obtain estimates of target distances ranging from 15 to 105 feet. Participants viewed a target for 10 seconds from a stationary position. As they viewed the target, they were asked to form a mental image of the target and its exact location. They were then blindfolded and asked to walk to the target's location, keeping the target's location in their minds as they approached it and stopping when they thought they had reached it. They were asked not to count steps or time in their heads. The distance judgments were performed both in a real world officer corridor and in a virtual office corridor modeled to simulate the real world corridor. The target, a construction cone, was clearly visible and distinct from the background at all distances. In general this method produces more accurate distance judgments than any of the other methods and is less prone to measurement errors.

Dismounted Warrior Network Research. Lockheed Martin Information Systems (1997) evaluated the ability of eight participants to perform a series of infantry tasks using several different VE interface configurations. Among the tasks evaluated was estimating the distance to enemy personnel located 164, 328, 492, or 820 feet from the observer in desert terrain. The targets were presented 32 times at each distance to each participant. The estimates were made in conjunction with target acquisition task and required the participants to verbally report the distance to the targets. This method is similar to the magnitude estimation procedure except the estimates are not made relative to a standard.

Findings

VEPAB Experiments. Experiments conducted with VEPAB indicated that distance estimates were less accurate and more variable in VE than in the real world, and were affected by

display characteristics. Lampton et al. (1994) found that estimates of distances to a human figure in VE were less accurate and had higher variance than estimates of distances in a comparable real-world task. Real-world distance estimation was inaccurate for the initial condition in which distance was to be estimated for a human figure of known height. However, after being given feedback for the first estimate (the figure was 40 ft. away), the real-world distance estimates were accurate as the figure moved toward the observer. VE distance estimation did not improve at shorter distances, contrary to expectations that stereoscopic depth cues should assist range estimation at shorter distances.

Lampton, McDonald, & Singer (1995) and Lampton, Gildea, McDonald, & Kolasinski (1996) found differences among display devices when estimating distances in a virtual environment. They also found that estimates made in the VE were significantly different from real world estimates using the same task and procedures. The monitor produced the least accurate distance estimates. In all cases the accuracy of distance estimates, as reflected by the larger relative errors, tended to decrease as the target distance decreased (See figure 2).

Singer, Ehrlich, Cinq-Mars, & Papin, (1995) found that both head coupling and a stereoscopic view significantly improved the accuracy of estimates at distances shorter than 10 feet.

Projective Convergence Experiments. The primary finding in our early studies of configuration learning (Witmer et al, 1995, 1996; Bailey & Witmer, 1995) was that our participants performed very poorly. They made gross errors in judging both the distance and direction to targets not currently in their line of sight. For example in one experiment, the average error difference between the actual target location and the judged target location was 86 feet. One explanation for these findings is that VE degrades distance estimation to the degree that it adversely impacts learning spatial layouts. Another possibility is that our participants became disoriented in the VE and were mostly guessing because they had no idea where they were or where the designated target was located relative to their current position. Two experiments that investigated configuration learning in outside terrain found differences in measures of configuration learning as a function only of the particular terrain database used, with a database with more distinctive terrain features producing more configuration learning that one with less distinctive features (Singer et al, 1997; Singer et al, 1998). In a follow-up analysis of the distance estimates provided in Singer et al. (1997), Allen & Singer (1997) found overestimation of distance for near landmarks and underestimation of distances for far landmarks using distinctive terrain, but no distance effects when using non-distinctive terrain. The type of VE (High or Low) used had no effects on the distance estimates. The High VE Group moved through the VE by walking on a treadmill and could scan the VE simply by turning their heads (head tracking). The Low VE Group was seated and used a joystick to both move forward and to look around.

Magnitude Estimation Experiments. Our previous work had shown that participants often bumped into walls when navigating in a virtual building. This suggests that they did not accurately perceive depth as they approached a wall. The magnitude estimation experiments identified a number of factors that influenced the accuracy of distance estimates. Kline and

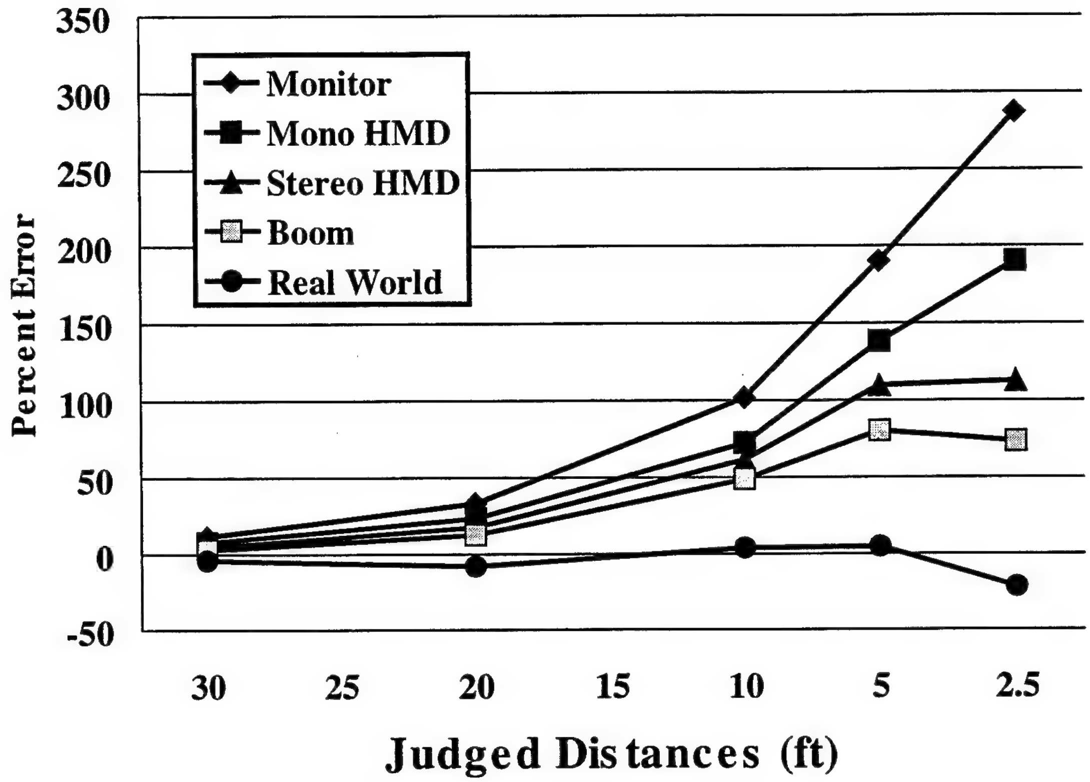

Figure 2. Relative error in distance estimates as a function of display device.

Witmer (1996) conducted a study to investigate how accurately stationary observers could estimate distance to a wall in a VE as FOV, texture, and pattern were varied. The observer's view was fixed in that head tracking was turned off. The distances being judged were between 1 and 12 feet. The results indicated that a wider FOV produced more accurate estimates than a narrow FOV. Distances were typically underestimated with the wide FOV and overestimated using the narrow FOV. A target placed 5 feet from the observer was judged to be at 2.68 feet with the wide FOV and 8.73 feet with the narrow FOV. Significant two-way interactions of distance with texture, pattern and FOV indicated that these variables affected depth perception only at the shorter distances.

In another experiment, Witmer & Kline (1998) investigated the effects of texture and pattern on distance judgements over longer distances, ranging up to 110 feet. The observers were stationary and had a fixed view of the target scene (i.e., no head tracking). Participants grossly underestimated the distance to the targets; the estimates averaged about 50% of the true target distance. This compares to estimates of approximately 75% of the true distance in a comparable real world environment. Cylinder size, distance, and the interaction of cylinder size and distance significantly affected the magnitude of the VE estimates. The estimates were more accurate for the small cylinder than for the large cylinder. For example, a target placed 50 feet from the observer was judged to be 22.57 feet for the small cylinder and 18.91 feet for the large cylinder.

Texture did not significantly affect either the distance estimates or the magnitude of the relative errors.

Witmer & Kline (1998) also reported the results of an experiment in which moving observers judged distance traversed for distances up to 280 feet. Half of the participants received compensatory cues (an audible tone every 10 feet) to help them calibrate their distance judgements to the true target distances. Although these cues were provided on only half of the trials, they improved performance to levels approaching perfect performance. The judgments averaged 96% of the true target distance when compensatory cues were present but only 67% of the target distance when compensatory cues were absent. The mode of locomotion used in moving through the VE (treadmill, joystick, or teleport) did not significantly influence the accuracy of the distance estimates, but speed of movement had a significant impact on estimation accuracy. Distance judgments were more accurate at the slow speed than at the fast speed. For example, targets at 280 feet were judged to be 267 feet on the average when moving at the slow speed and 241 feet when moving at the fast speed. Accuracy of the distance estimates generally decreased as distance to the target increased.

Non-visually Guided Locomotion Experiment. The extremely poor VE distance estimates made by a stationary observer and the lack of substantial improvement in the accuracy of the estimates when observer movement is added (Witmer & Kline, 1998) suggest that either verbal estimates of distance are not very accurate or that VEs degrade distance estimation to a large degree. To determine how much of the problem is due to the requirement to provide verbal estimates of distance and how much is due to VE factors, Witmer & Sadowski (in press) used NVGL to obtain distance judgements in VE and real world environments. The judgements were made in an office building corridor with targets placed at distances between 15 and 105 feet. The distance judgements averaged about 85% of the true target distance in the VE and 92% of the true target distance in the real world environment. The differences between the distance judgements in the VE and in the real world were significant, however. The magnitude of the errors in the VE was nearly twice those obtained in the real world. The order in which the environments were presented also produced a significant result. Distance estimates were more accurate when the real world environment was presented first. The beneficial effects of providing feedback were not statistically significant in this experiment (p=.06), but the data suggest that more trials with feedback could be beneficial.

Dismounted Warrior Network Experiment. Lockheed Martin (1997) reported that verbal estimates of target distance were very inaccurate at target ranges of 164 to 820 feet. Our analysis of their data revealed that the estimated distances averaged 234% of the true target distances. These gross overestimates of target distance contrast sharply with what we have typically found. The practically featureless desert database used in this research coupled with a very small target size is likely responsible for these unusual results. Pleban, Dyer, Salter, and Brown (1997) reported that the desert terrain made it difficult if not impossible to estimate distance, and that an enemy target at 407 feet appeared as but a speck on the horizon in the VE displays.

Discussion

Our early investigations of configuration learning suggested that distance estimates in VE were poor. A second line of research, comparing VE and real world performance on the VEPAB distance estimation task, showed VE estimates to be significantly less accurate than real world performance. This suggests that the poor performance in VE is at least partly due to perceptual difficulties that affect the accuracy of distance judgments in VE. Follow-up work by Lampton et al. (1996) showed that the type of display device used influenced the accuracy of distance estimates in VE. The primary differences between the devices investigated were their FOV and their resolution. Subsequent work by Kline & Witmer (1996) demonstrated that reducing the FOV for one of the devices (BOOM2C) could affect not only the amount of error in distance estimates, but also the direction of that error (underestimates vs. overestimates). The narrow FOV produced less accurate estimates by reducing or eliminating linear perspective cues. Surdick, Davis, King and Hodges (1997) found that perspective cues (linear perspective, foreshortening, and texture gradient) were more effective for judging distances up to about seven feet than were other depth cues, including relative size, relative height, and relative brightness. Singer et al. (1995) demonstrated that distance estimation accuracy at distances shorter than 10 feet was improved by the addition of head-tracked stereoscopic displays. Witmer & Kline (1998) reinforced Lampton's et al. (1996) conclusion that the accuracy of distance estimation in VEs is less than in real world environments. Furthermore, they found that manipulation of textures did little to eliminate the observed deficits in performance. Although target size did influence performance, manipulation of the size of unfamiliar objects is not a practical solution. Taken together, these studies suggest that VEs distort monocular or stereoscopic distance cues, negatively impacting the distance judgements in those VEs.

We hoped that providing the cues for distance associated with movement through the VE would compensate for the distortion of other distance cues, resulting in substantial improvements in performance. However, Witmer & Kline (1998) found that neither movement method nor edge rate markedly changed the distance judgments. These results indicate that proprioceptive cues and visual flow cues may not play a major role in making distance judgements in a VE. In contrast, movement speed did influence distance judgments, suggesting that the time to cover a distance changes one's perception of distance traveled. This research also suggested that distance perception in VE could be recalibrated cognitively by providing compensatory cues for distance. This cognitive recalibration may be only temporary, however. Witmer & Kline (1998) did not collect any data that would answer questions about transfer of estimating skill to other distances or environments.

Using NVGL to evaluate the accuracy of VE distance estimates altered our working hypothesis regarding how much VE degrades distance estimates. This procedure yielded more accurate VE distance estimates, suggesting that the use of verbal distance estimates is partly responsible for the poor performance observed in our research. However the magnitude of the errors in VE was still twice that observed in the real world, establishing beyond a shadow of a doubt that the VEs are distorting perceptual judgments of distance.

What factors might be responsible for this distortion? In our search for an explanation it is important to remember that the performance decrements were found across various VEs using different display devices, and with varying movement conditions. It is also important to keep in mind the distances investigated in each experiment, because the effective range of various distance cues vary with the distance being judged.

To understand why VE tends to distort distance perception at the target distances investigated, we need to know which distance cues are effective at those distances, and to assess the extent to which these cues were present or absent in our research. Cutting and Vishton (1995) have identified which depth cues are most effective at different distances and related these cues to three egocentric regions or zones of space: (1) personal space extends just beyond arms reach and refers to space used by a static observer; (2) action space extends to about 100 feet and refers and includes distances in which an observer can throw an object to another person or easily talk to others; and (3) vista space extends beyond 100 feet. Research employing VEPAB and Kline and Witmer (1996) studied both personal and action space. In personal space the most important depth cues are occlusion, binocular disparity, relative size, convergence and accommodation. The remaining studies investigated action space and vista space. The primary distance cues in action space and vista space are the pictorial cues, including occlusion, height in the visual field, convergent linear perspective, relative size, and relative textural density. In addition, two other distance cues, binocular disparity and motion perspective are effective distance cues in action space. Note that accommodation and convergence are not effective depth cues in action space or vista space.

Witmer and Kline (1995) have shown that while relative textural density influences distance estimates in VE, its effects are typically too small to account for the differences between real world and VE distance estimation performance. Similarly adding observer movement, which provides motion perspective and other movement related cues does not eliminate the deficits in performance in VEs (Witmer and Kline, 1998). Research by Wright (1995) and Witmer and Kline (1996) suggests that simply using a high resolution or wide FOV VE display cannot erase the deficits in perceived distance. Although occlusion is probably the most powerful depth cue in action space, it was not a factor in our distance estimation tasks. Of the remaining distance cues listed by Cutting and Vishton (1995), height in the visual field, convergent linear perspective, relative size, and binocular disparity appear to be the most likely candidates for explaining the observed discrepancies between VE and real world judgements of distance.

The National Research Council (1997) has suggested that the restricted FOV provided by VE displays must degrade height in the visual field and convergent linear perspective as cues for distance at some point. Relative size as a distance cue may be degraded as well. The limited vertical FOV found in most VE displays (ranging from 40 to 90 degrees) may be responsible for this degradation. By comparison, the real world vertical FOV is approximately 120 degrees. A reduced vertical FOV may result in distant objects appearing closer in VE than they would in the real world because these objects would be compressed into a smaller visual frame as they recede into the distance. Kline and Witmer (1996) showed that a reduced horizontal FOV could also adversely impact the accuracy of distance estimates by reducing or eliminating linear perspective cues. Because linear perspective cues are among the most effective distance cues in simulated

environments (Surdick et al., 1997), reducing or eliminating these cues can have a major impact on the accuracy of distance estimates.

In VEs, emulation of binocular disparity is achieved by presenting different images to the two eyes with some central area overlap. While this technique may provide the illusion of depth in VE, it may not faithfully reproduce real world depth. Cutting and Vishton (1995) noted that early stereoscopic pictures enhanced the distance between the eyes to show large expanses and cityscapes, diminishing the effective size of the objects seen. Relative size may be important factor at the closer distances because the perceived size of an object accelerates as the distance to the object decreases, yielding a looming effect. Accommodation and convergence cues are not accurate in VEs, a fact that researchers often use to explain poor distance estimation in VEs. However, these cues are only important for judgments in personal space and at the shorter distances within action space.

Another remote possibility is that the display field of view and geometric field of view (GFOV) were not properly matched. The GFOV is the FOV of the image displayed by the image generation software on the display device. If the GFOV and the display device FOV match, the image appears "normal." A mismatch will result in minification or magnification of the image sizes. To obtain the consistent pattern of underestimation of distance in VE, the images would need to be magnified, causing objects to appear closer to the observer than they actually were. This could only occur if the GFOV was set smaller than the display FOV - an unlikely scenario given the small FOV for most of our displays. Some studies (Kline & Witmer, 1996; Witmer & Kline, 1998) made rough measurements of the display FOV to ensure that the GFOV was set to match the display FOV.

Additional research is needed to determine which of the distance cues operating in action space are most responsible for degrading distance judgements in VE. Once the causes of this degradation are isolated, we can begin working toward a solution. The solution may be as simple as increasing the VE display vertical or horizontal FOV, or adjusting the overlap in VE stereoscopic viewing devices. On the other hand, it may involve major technological advances, such as inventing new techniques for emulating binocular disparity in VE displays.

Interface Effects on Performance

Display Device Effects

Resolution. The VEPAB tasks include a VE Snellen eye chart and VE color vision tests. Lampton, Gildea, McDonald, and Kolasinski (1996) examined performance on these tasks by participants using an HMD, a Binocular Omni-Orientation Monitor (BOOM), or a standard computer monitor.

For acuity determined by the Snellen eye chart, scores were ordered according to estimates of resolution based on the horizontal pixel density of the displays, but were worse than what would be predicted from the resolution estimates. The mean Snellen acuity score was 20/500 for the HMD, 20/200 for the BOOM, and 20/91 for the standard computer monitor.

Acuity varied across participants within each display group even though they had been screened for normal vision in the real world.

Color. Lampton et al. (1996) also found that some participants who passed the real-world color vision test did not pass the VE test, and vice versa. This may have resulted from imperfections in the method in which the color test plates were digitized through scanning, errors in software rendering, imperfections in the display devices, or an interaction of these factors. Boff and Lincoln (1988) noted that the luminance and flicker of a display can affect color appearance. Inconsistent representations of color in VEs have implications for military training applications; for example, variation in color presentation across devices or within a device across sessions may result in training exercises in which the detectability of camouflaged troops or vehicles varies greatly and inappropriately for training purposes.

Stereoscopic Viewing. As reported earlier, Singer et al. (1995) found that stereoscopic viewing improved distance estimation at target distances of less than ten feet. Conversely, it also increased the number of collisions with walls and doorways on the initial trials in a locomotion task. With practice, this effect disappeared.

Self-representation

Anecdotal reports, observation of participants in our experiments, and our own experience in VE all suggest that one potential obstacle to successful task performance and learning, particularly involving the acquisition of spatial knowledge, is the loss of egocentric orientation that can occur in VE. (Egocentric orientation is understanding spatial arrangements including objects or places in relation to self. Exocentric orientation refers to spatial understanding of objects in relation to one another.) If the user is seated and using a joystick for movement control, or walking on a fixed-direction treadmill, vestibular feedback will not always agree with visual cues for movement. This mismatched cue set may impair performance, distort learning of the task, or detract from the transfer of task or situation context to the real world. Another related factor is the nature of physical interaction required by the task. Interaction with the simulation that is coupled to appropriately spaced, body-linked hand movement should provide more egocentric orientation and proprioceptive feedback than a disembodied interaction (e.g. a cursor or generically modeled hand fixed in the visual display, a joystick maneuvering a cursor, etc.). It is possible that the low physical and functional fidelity of such representations will result in decreased learning, retention, and/or impaired transfer.

There is little research on the psychological and performance issues in body representation in VE systems. A key issue is the influence of body representation as an anchor for experience and possibly an egocentric reference point for interaction. Draper (1996; Draper, Wells, Gawron, and Furness, 1996) has done a series of experiments on the effects of a virtual body representation on the performance of object positioning and reaching tasks. In two experiments, having a body model (arm and hand) did not influence the amount of error that subjects made in re-positioning target objects. In the third experiment (Draper, et al., 1996), the errors in positioning in order to reach a target (the experimental task) interacted with the degree of body representation; a disconnected hand representation led to more errors at low or medium

heights and the full body representation created more errors at the higher height. At this higher height, the full body representation resulted in a subject's viewing a reaching arm against a low texture background (ceiling), which may have combined to distort the subject's perceptions for reaching. The lower levels provided more texture (floor and furniture) and more of the body representation, which may have added cues for the reaching task that overcame misleading cues based on the arm representation alone. In that third experiment, target height was a main effect, which may have been solely due to the amount of texture available at each target height. It should be noted that the number of subjects used in each experiment was very low (less than ten per group).

Singer, Ehrlich, & Allen (1998) investigated the influence of a virtual body model in VE on performance during a simple search task in an office environment. The hypothesis was that the presence of a body model would result in better awareness of body orientation and position, and therefore better movement in the VE. Several sets of rooms were developed in which participants could search for a simple target, a briefcase. The body model participants experienced the VE with a body representation linked to their point of view, shoulders, arm, and hand. Sensors were also attached at the ankles to monitor relative limb position. When one foot was raised, the corresponding foot on the body representation was raised, and a step (average stride length) was taken in the VE. The non-body model group experienced the same VE with a head-coupled point of view, but only a wand pointer representing their hand position. Measurements included time for searching, time to acquire targets, and time to exit target rooms. In addition, the overall number of targets acquired and the collisions experienced during those phases were also analyzed.

There were no significant differences in times or collisions between the body representation conditions. Participants improved in both speed and number of collisions in both. Finding the same pattern in the speed and the accuracy measurements for both body representation conditions indicates that self-representation (having a body) did not aid in the adaptation to movement through this VE configuration. The VE interface was set up to allow direct locomotion control, which provided reasonably good kinesthetic cues. Orientation in the real world and in the VE was identical. This may have supported the improvement during the movement-oriented segments of the trials.

The restricted FOV of the HMD (48 degrees horizontal by 36 degrees vertical) may have contributed to the difficulty of the search phase of the task. The importance of the size of the visual field has been demonstrated in research on performance of guiding-movement tasks. For example, Wood and Troutbeck (1991) investigated the effect of FOV on driving performance. They found that restricting the visual field from "normal" (approximately 120 degrees for the binocular field; Kaufman & Christiansen, 1984) to forty degrees, and even further to twenty degrees, decreased peripheral awareness, impaired obstacle avoidance, increased maneuvering errors, and increased time to complete the course. Impaired obstacle avoidance during driving is comparable to the search phase in this experiment, as obstacles have to be detected before they can be avoided. Following this logic, it seems reasonable to predict that increasing the FOV would ease any visual search activity, decreasing the time for visual search and therefore decreasing the overall time required for any "search and acquire" task.

The results of this and other experiments (e. g. Draper, 1996; Draper, Wells, Gawron, and Furness, 1996) on body representation seem to indicate that body representation is not as important factor for movement in VE as might have been expected. These results should not be taken as definitive. The results may not indicate the need for body representation, but the need for some interacting mix of FOV, visual depth cues, body representation, and task parameters.

III. Training Skills in Virtual Environments

When the training portion of our research program was initiated, we decided to investigate first those tasks that VE systems appeared to be most capable of training, spatial learning tasks (e.g., Witmer et al, 1995; Singer, Allen, McDonald, & Gildea, 1997). Aspects of some of these experiments will be discussed below in order to explain what we have learned about using VE technology to train tasks with the goal of transferring the skills acquired to the real world.

Skill Acquisition and Transfer

There are several interacting factors required to use VE systems effectively for training or rehearsal. Foremost among these are the fidelity issues that have been repeatedly addressed in simulation-based training (Hays & Singer, 1988). The first issue is the degree to which the system supports the required task interactions. In other words, the cues, response capabilities, and feedback have to be sufficient for minimal task performance. If the trainee is unable to perform the task, and therefore unable to improve on that performance, learning may be minimal. Secondly, correct actions or errors made in the VE must result from the same fundamental human processes as in the real world. Finally, the VE configuration should present stimuli and support actions that are sufficiently like the real world to transfer easily to the intended real-world scenarios. When these cue, response, and feedback conditions are satisfied, learning can be demonstrated through task performance improvement in the VE and transfer can be tested in the real world. Transfer to the real world is the goal of simulation-based training (e.g., Hays & Singer, 1988).

VE technology provides new methods and techniques for replicating equipment, task parameters, and environmental conditions that were previously difficult or impossible to simulate. VE approaches do this by providing more realistic representations of the physical characteristics and functional aspects of the to-be-learned situation. VE also allows the addition of many aiding, augmenting, and adjunct features that can serve to speed learning, improve retention, and/or increase transfer. Unfortunately, each VE configuration also may have interface characteristics that bias the user's experience, possibly to a sufficient degree that transfer is slowed or impaired. These interface characteristics generally reflect either the unavailability or the high cost of the technology required to replicate reality exactly. Examples include limited visual display resolution and FOV, system lags in responding to trainee actions, and low-fidelity simulation self-motion. As a result, system designers develop "work-around" solutions, such as the use of a joystick to control locomotion, or the use of enlargement or color to make distant objects visible. For example, much of the training and rehearsal for dismounted soldiers will require controlled movement through simulated areas (e.g., urban terrain). Therefore, simulations of walking will be background context for these tasks. However, current walking simulations for controlling movement in VE require physical movements and provide feedback which differ from walking in the real world to a greater or lesser degree (e.g., Singer, et al., 1997). They therefore require at least some minimal level of adaptation. Similarly, "clicking" on an object with a hand-held controller differs from actually grasping it and picking it up.

The effects of these work-arounds on training and transfer is a matter of concern, because of both the potential need to acquire new, VE-specific skills in order to train in VE, and the potential effect on the transfer of the new skills to the real world. Learning some VE-specific skills is a matter of generalizing from a pre-existing, well-learned skill so that the altered performance is conducted in the presence of a new set of stimuli. For example, we instrumented a treadmill that enabled participants to walk through environments at normal speeds (Singer, et al., 1997; Witmer & Kline, 1998). Walking on the treadmill required pushing a wide belt toward the rear of the apparatus, which rotated an instrumented roller, allowing registration of movement in the VE. Participants had handrails to hold (and lean against), and buttons on each handrail to control orientation for movement. Our participants readily adapted to the VE, and quickly readapted to movement in the real world (they could walk normally after the experiment). Walking is an example of an over-learned and flexible skill. Walking in the VE required adapting the well-learned skill of walking in the real world in the presence of VE stimuli (HMD, Treadmill). It is unlikely that any walking changes required for interaction in the VE would detrimentally transfer to real world activities after the experience.

The ease of adaptation to the VE simulation does not carry a guarantee that there will or will not be detrimental transfer or interfering real-world re-adaptation effects. What may be more important is the degree of generalization and discrimination from preexisting movement skills and the newly acquired VE movement skill. If the to-be-learned skill is not easily generalized from a preexisting skill, then a new skill must be built out of some combination of the existing repertoire of skills and some new basic skill. As an example, if someone is being trained to identify possible threat areas based on potential fields of fire, many existing visual skills generalize to orienting, observing, and identifying objects and locations in the VE. If the HMD restricts vision, some new scanning technique may be required for interaction only within the VE. Finally, a new cognitive appraisal skill for projecting a potential field of fire based on a single visual inspection would need to be learned. As a result, in the training scenario, the learner is using normal visual skills, combined with a new VE-specific scanning skill, in order to learn (for ultimate transfer to the real world) the cognitive-spatial skill of threat appraisal.

The first team training experiment in our program (Lampton, McDonald, Rodriguez, Knerr, & Parsons, 1998) requires participants to aim and fire a simulated pistol. Sensors on the elbow and on a manual control device allow the participant to aim the pistol. In the real-world, our sense of joint position, and other non-visual perceptual systems, allow us to shoot at a target without the pistol being in our FOV (shooting from the hip). However, in VE, imperfections in the sensor system sometimes result in a mismatch between the felt position and the representation of the hand (or pistol) in the VE. To compensate for this discrepancy, a laser beam projecting from the pistol is represented to help the participant aim. Preliminary testing indicates that participants quickly learn to use the manual control device.

There are two areas of concern regarding the transfer of skills from VE to the real world. England (1995) proposed that because of inadequate feedback in VE, practice in VE of a skill learned in the real world could become "de-skilled" with respect to the ability to use the skill in the real world. He pointed out that although VE users may be able to adapt to unnatural interaction in VE, the introduction of stress may lead them to unconsciously revert to earlier

learned (real-world) interactive skills that are inappropriate for VE interaction. He stated that for real-world sensory-motor manipulation tasks, sources of feedback include cutaneous and kinesthetic receptors, vestibular apparatus, vision, and audition. These sources of information overlap and offer redundancy, decreasing the uncertainty of perceptual cues. In contrast, the lack of such redundancy in VE may lead to changes in the performance of manipulation tasks. In VE, cues are primarily visual, and even those may lack the precision and reliably of real-world visual cues. While England has built a strong logical case, there is no direct empirical evidence that it has actually happened. This possibility seems more likely with highly skilled, infrequently practiced motor skills than with tasks such as walking.

The other concern is that practice in the VE provides extra skill or knowledge that is either required for performance or improves performance in the VE, but inhibits transfer. Some consideration must be taken in the design of any simulation to ensure that those detrimental incidental skills, behaviors, and/or non-real cues are not "attached" to the to-be-transferred skill.

Psychomotor Skills

Locomotion

In most dismounted soldier training applications, trainees will need to move at realistic rates while performing other tasks. Certainly, users will need to be able to move in a VE in a way that does not detract from or interfere with the performance or learning that is the goal of the VE session. In our experiments, we found that participants were able to move through the VE successfully with a variety of control devices. Some resulted in better locomotion performance than others. Other researchers have identified problems with devices that move the user in the direction of gaze. The evidence suggests that a walking simulator may produce better spatial learning than hand or eye control of movement.

The VEPAB contains a series of locomotion tests that vary in difficulty and can be used to train or evaluate movement in VE systems (Lampton, et al., 1994). Experiments in spatial learning have used button control (Witmer, et al., 1996), joysticks (Witmer, et al., 1994), an instrumented treadmill (Singer, et al., 1997), and a "walking simulator" (Grant and Magee, 1997) for movement control. Research in distance estimation has used joystick and treadmill for control of movement (Witmer & Kline, 1998; Witmer & Sadowski, in press). In some cases the control of movement was incidental, while in other experiments it was an independent variable (see Table 3). Basic to these (and many) VE tasks is the need to move at reasonable speeds and avoid collisions with virtual objects by using an interface to control the direction and rate of movement of the user's viewpoint.

The first two VEPAB experiments (Lampton, Knerr, et al., 1995) investigated movement through the VE using either a joystick or a spaceball for movement control. The joystick was a standard two-dimension device that used two buttons for additional input (e.g., press a button and push the joystick forward to fly up). The spaceball was a six degree-of-freedom device with minimal physical movement that used pressure input to control movement (e.g., press forward

Table 3

Summary of Experiments Using Different Movement Control Interfaces

Exp. ID	Interface Type	Status	Effect
1.1	Joystick, Spaceball	Comparison	Joystick better Practice effects for some tasks
1.3	Joystick, Spaceball	Comparison	Joystick better Practice effects for all tasks
1.4.	Joystick	Constant	Practice effect for locomotion task.
2.1	BOOM, Real World	Constant	BOOM practice trials took longer than real world practice trials
2.2.	Joystick	Constant	Practice effects
3.1.	Joystick (Low), Treadmill (Hi)	Compared (as part of VE Condition)	No Significant Differences
3.2.	Joystick	constant	N/A
4.2.	Joystick, Treadmill, Teleport	Compared as cues for distance estimation	No significant differences
4.3	Treadmill, Real World	Compared	RW > Treadmill
5.1	Mime-walking	Constant	Practice effect

and pull up to fly forward and up). There were seven locomotion tasks in the VEPAB suite of tasks:

Straight-away required moving down a corridor to a target, rotating 180 and returning to the starting position,

Backup repeated the straight-away task to the target, but returned by backing up,

Turns had ten 90° turns in a narrow hall with varying lengths in the straight segments.

Figure-8 was a continuous course of two unequal sized oval corridors that crossed.

Doorways consisted of a series of ten rooms connected by doorways.

Windows was a flying task that required three-dimensional movement through ten rooms connected by door-sized windows.

Elevators was a flying task that required horizontal and vertical movement while going over and under vertical partitions.

The experiments found a significant difference between the joystick and spaceball for the movement tasks, with the spaceball requiring roughly twice the time for performance that the joystick required (Lampton, Knerr, et al., 1994). Significant practice effects were found for the first locomotion task performed (straight-away) and the first locomotion task that involved flying

(windows) in experiment 1.1. In experiment 1.3, which involved 11 times as much practice, significant practice effects were found for all three locomotion tasks used.

Witmer & Kline (1998) found that walking at a guided pace on a treadmill did not now produce significantly different estimates of distance traveled than did using a joystick set at a constant speed. Their explanation is that the treadmill walkers had to perform a secondary task, maintaining a constant pace, which interfered with their primary task of distance estimation (See page 12.) The proprioceptive cues that the treadmill interface provided did enhance the sense of presence in the VE.

Witmer and Sadowski (in press) found blindfolded walking to a previously seen target in the real world to be more accurate and less variable than in the VE. (See page 12.) They attributed the differences to visual display inadequacies, not psychomotor differences, although the participants walked more slowly (and hence took longer for the same distances) in the VE condition.

One experiment on spatial learning in building interiors used the buttons on the BOOM2C (Witmer, et al., 1996) for movement control. In the experiment, the buttons on the BOOM2C moved the user forward or backward along the line of sight, while the display was held to the face (the display and buttons are mounted on the end of a counterbalanced arm). The subjects in the experiment practiced in the VE before transferring to the real world building. Practicing the route in the VE took longer than did practicing the route in the actual building, although the difference decreased substantially over the course of three practice trials.

Two experiments investigating spatial learning in open terrain (Singer, Allen, McDonald, & Gildea, 1997; Singer, Allen, McDonald, & Fober, 1998) used a treadmill or a joystick for movement control. In the first experiment, the treadmill and the joystick were used as the basis for investigating the effects of different VE configurations on spatial learning. In the second experiment, only the joystick was used to control movement through the VE. The time and speed of interaction were controlled in these experiments, and differences in spatial learning and orientation in the large scale environments (several kilometers on a side) did not appear to be strongly related to movement control (Singer, et al., 1997; Singer, et al., in preparation). Further analyses of distance estimation and landmark identification during the training phase of the first experiment found differences that seemed to be the result of increased cognitive and psychomotor load on the High VE (treadmill and head-tracked) group. The treadmill using group seemed to require more thought as well as more exertion in controlling movement through the environment, which in turn may have affected the acquisition and use of proprioceptive cues to distance and direction (Allen, McDonald, & Singer, 1997).

The treadmill problems, requiring greater exertion with smaller strides (and button control for direction orientation), have been addressed by the development of sensor-controlled walking techniques. The first test of this technique was an experiment investigating the effect of body representation on search task performance (Singer, Ehrlich, & Allen, 1998). (See pages 17 and 18). Sensors were attached at the ankles to monitor relative leg position, and a sensor on a shoulder harness monitored body orientation in space. When one foot was raised, the

corresponding foot on the body representation was raised, and a step (average stride length) in the environment was taken. No comparison with other modes of movement were made in this experiment, but the subjects adapted quickly to the technique, and movement times for the search task decreased significantly over the course of the experimental trials.

Grant and Magee (1997) used a walking interface functionally very similar to that used by Singer, Ehrlich, & Allen (in preparation) in an experiment investigating the acquisition and transfer of spatial knowledge. The participants' task was to learn the location of a number of landmark objects in a science museum. During a training session, the walking simulator and joystick groups made similar orientation errors when asked to point to their original starting location. During a transfer test in the actual building, the walking simulator group walked significantly less distance to find all of the landmarks than did the joystick trained group, but they did not find them significantly faster. Grant and Magee (1997) claim this as evidence that there is something learned from the walking interface *in addition to* the visual information that can be used for spatial navigation.

The DWN experiments (Lockheed Martin, 1998) compared four different interface suites in terms of their ability to negotiate a fixed course thorough a simulated built-up area, including building interiors. Maximum possible movement rates were the same for all systems. An interface (VIC A) which used a foot pedal to control forward and backward movement in the direction of gaze produced both more rapid course completion and more collisions with walls and other objects than did the other interfaces: a human joystick (automatic forward acceleration and movement when standing in a pre-defined area), a sophisticated joystick, and an omni-directional treadmill. The authors attributed this result to a combination of two factors: the foot pedal encouraged full-speed movement, while the gaze-directed steering produced movement toward whatever object the soldier was looking at (perhaps in order to avoid it). Other research on movement control in constrained environments has found gaze-directed steering to be no better than hand-pointing when moving directly to a target, but found hand-pointing better when moving relative to targets (Bowman et al., 1997). This seems to support Lockheed Martin's collision avoidance problem hypothesis in the gaze-directed foot-pedal movement control interface. In addition, soldier participants reported that it was more difficult to move in a straight line, maintain balance while moving, and change direction while moving with the omni-directional treadmill than with the other interfaces

The omni-directional treadmill is one of several physically controlling techniques (walking surrogates involving leg and foot movement) for controlling movement in VEs that have been developed. Templeman (1997) is developing techniques similar to those developed by the Institute for Simulation and Training (IST) for Singer et al., (1998), that use sensors on the legs and feet to monitor limb position and control representative walking in the VE. In Templeman's system, pressure sensors on the feet interact with sensors just above the knee, providing input to pattern recognition software. The technique allows control of both forward and backward movements, as well as sidestepping. When the knee moves forward and back (as it would when walking in place) a step is taken forward. Moving the knee back and then forward allows backing up, lifting out to the right initiates a rightward side-step, and so on. The size of the virtual step is a function of how far the sensor moves, and how rapidly motions are made. As

Templeman (1997) points out, and as with the walking system developed by IST, the walking motions made using these artificial systems do not use exactly the same motor routines that are used in normal movement.

Taken together, these results seem to suggest that there are some advantages to a "walking simulator" over joysticks, gaze-directed, or hand-pointing controls in terms of training effectiveness. However, the cause of this advantage is not clear. Is it because it provides better stimulus-response compatibility between the training and transfer conditions, better cues for changes in direction, better cues for distance traveled, or simply because it places less cognitive demand on the trainees, allowing them to attend more carefully to other cues? The only drawback anticipated is that the walking movements are not exactly the same in terms of normal neuromuscular activities. However, as pointed out above, it is unlikely that any adaptations to excessively well-learned routines like walking would detrimentally transfer to real-world activities. The great benefit for dismounted soldiers is that simulated equipment (e.g., a rifle) can be carried and used instead of having to control movement with hand controls.

Manipulation

Research in our program has only investigated manipulation tasks in limited ways. Most work has been done with representative tasks that primarily involve selecting and moving ("dragging") objects from one position to another, or just pointing and selecting. Overall, participants can learn to perform these actions successfully, but precision and speed of movement are limited. The experiments using VEPAB provide some information about how well people can use different interfaces to learn relatively simple manipulation tasks. The Bins task required manipulating a 3D cursor to acquire a ball in one bin and maneuver it to another, targeted bin. Other manipulation tasks included using the cursor to rotate a dial and move a slide control. These were all performed significantly better using a joystick interface than a spaceball (Lampton, Knerr, et al., 1994). The interface made no difference in Tracking tasks. However, this experiment was conducted with a system with a low update rate, which made the tracking task very difficult with any interface device. Several of the VEPAB tasks were also used in the investigation of visual presentation differences (Singer, Ehrlich, et al., 1995; Lampton, Gildea, et al., 1996). In these experiments, participants easily learned to use the joystick interface in the context of the simple tasks. Practice effects were usually found.

We have since used a sensor positioned on the hand to orient a virtual vector (a colored line) and point at landmarks (Singer, Allen, et al., 1997) or at targets in a room (Singer, Ehrlich, et al., 1998). In spite of the slight offsets between kinesthetically sensed hand position and visually represented hand and vector orientation, participants were able to adapt to pointing almost immediately. The accuracy of pointing was moderately good, especially at the short ranges in the search experiment (approximately one meter).

Kozak, Hancock, Arthur, and Chrysler (1992) reported a failure to demonstrate transfer of training on a motor task practiced in VE and tested in the real world. England's (1995) concerns discussed earlier are applicable. The lack of redundant cues in VE (cutaneous, kinesthetic, vestibular, and auditory) may lead to changes in the performance of manipulation tasks.

If the task requires some reaction to environmental stimuli for adequate performance or for training, then the physical speed with which stimuli can be generated and changed in the VE and the accuracy and speed with which body parts can be tracked are important. Humans can respond with speeds that vary from a few to hundreds of milliseconds (Boff and Lincoln, 1988), using motions that can vary dramatically in speed, range, and complexity. The major problem in VE task representations that range from simple, whole hand movements (e.g., throwing or catching) to complex psychomotor manipulations (assembly of complex parts) is in the accurate position tracking of the critical body parts (Durlach and Mavor, 1995). Technological capabilities still lag human capabilities in this area. Moreover, some VE equipment is encumbering and hampers rapid or successful task performance. Some examples will illustrate these points.

Lampton, Gildea, McDonald, and Kolasinski (1996) found that performance of a 360-degree search task was slower with a head-tracked HMD than with a monitor controlled by a joystick. The difference was attributed to the weight and mass of the HMD, which slowed head movements.

During the DWN experiments (Lockheed Martin Information Systems, 1997), soldiers fired simulated rifles at targets from a standing position using different simulators. Average absolute error from the center of a target at a range of 100 meters was approximately one meter. Soldiers using two simulators in which they could physically assume the prone position also attempted to fire from that position. After 15 trials it was obvious this change in position had increased the transmitter-receiver distance of the weapons trackers to the extent that neither system could engage targets reliably.

Psychomotor Conclusions

Movement and manipulation will be the basis for training and rehearsal of dismounted soldier tasks in VE systems. Participants in our research have readily adapted to movement and manipulation using a joystick, and pointing with a hand-mounted sensor. Our work has also shown that the different methods of movement are essentially non-interacting with distance estimation (Witmer & Kline, 1998; Witmer and Sadowski, in press), although there are VE-based differences in distance estimation. The movement control interfaces are easy to adapt to and did not affect spatial orientation in our work (Singer et al., 1997; Singer et al., in preparation), although others (Grant and Magee, 1997) have found superior performance in spatial learning with a walking interface. An experiment using a movement and search task (Singer, Ehrlich, & Allen, in preparation) did not find any differences in psychomotor control with the presentation of a body representation (in terms of speed or collisions). The same experiment did anecdotally determine that the representative sensor-based walking is easy to learn, although it is somewhat unnatural.

In general, participants in our research have learned relatively rapidly to use whatever control interface was provided in order to perform assigned tasks. The caveat is that in some cases we have designed our interface systems to be easily usable in order to investigate other issues. Given that caveat, it must be remembered that movement and manipulation becomes more difficult as required position accuracy and/or tracking speed increases. If the movement or

found that the walking simulator group did better on the transfer task than did a group that had no experience, while the joystick group did not. Neither VE-trained group differed significantly from a map-trained group. The walking simulator group took more time to locate all of the landmarks, but did not walk significantly farther to visit them than did the group trained in the actual building. Both VE groups made larger errors when pointing to their starting point than did the group trained in the actual building.

In contrast to the route learning results, we found no support for transfer of configuration or survey knowledge from VE to the real world (Witmer et al., 1995, 1996). More recently, Wilson, Foreman & Tlauka (1997), using several different measures of configuration knowledge, were successful in demonstrating transfer of configuration knowledge from a simulated building environment to the real world. Wilson et al. (1997) attributed their success in demonstrating training transfer of configuration knowledge to the use of a relatively simple environment in comparison with our more complex environment.

Spatial Learning: Open Terrain.

We investigated spatial learning based on the level of VE interaction with large-scale terrain models, while performing simple cognitive terrain appreciation activities. The acquired spatial knowledge was tested in the same VE by requiring landmark identification and distance estimation from sites different from those experienced during training. The High-Level VE (Hi-VE) group walked on an instrumented treadmill while wearing a head-tracked helmet-mounted display and using a pointing wand for indicating directions or locations. The Low-level VE (Lo-VE) group used a joystick to control view/movement, while observing the terrain through the same HMD without head tracking (with a fixed view, while seated). They moved through the same simulated terrains and performed the same activities as the Hi-VE group. The control group (Map) performed the same activities symbolically while using topographical maps, with paced study replacing movement through terrain.

After the training session, participants' knowledge of the terrain was tested. The participants trained with more highly interactive VE experiences developed significantly better spatial knowledge than participants trained in comparable map exercises. This difference held over two different terrains. Analyses of measures of projective convergence, which combine direction and distance estimations to produce a measure of knowledge, revealed a significant superiority of the VE conditions over the Map practice condition in one terrain.

Transfer to the real world was investigated using one of these two terrains, which was modeled on a small area at Fort Benning, GA. A single VE training condition, using a head-tracked HMD and joystick controlled movement, replicated the training used in the first experiment. As before, map-based training was used as the control condition. The participants in this experiment were trained U.S. Army officers (Rangers and Basic Infantry Officer Course graduates), none of whom reported familiarity with the actual terrain location. The VE terrain model was minimally adjusted (due to time constraints) to attempt to account for road and drainage work that had been performed at the area. The participants were first tested in the VE, by indicating landmarks from new, untrained positions on the terrain. As in the first experiment

the VE-trained group performed significantly better than the map-trained group in the VE test. When transferred to the field, the map-trained group performed as well as the VE-trained group. Given the training background of the participants, it is not surprising that they were able to transfer map-study to the field. Map study has been shown to be highly effective (Thorndyke & Hayes-Roth, 1980). The interesting finding in this experiment is that the VE-trained group transferred approximately equivalent spatial knowledge to the field.

Cognitive Skill Transfer to Real World

Spatial learning is the only cognitive skill addressed in our research to date. The results indicate that people can learn about the landmarks (Singer, et al., 1997) and route learning (Witmer et al., 1995, 1996) in VEs with relatively short periods of exposure. We have not shown that VE produces better spatial learning than real world experience in these short sessions (Witmer, et al., 1995, 1996), but we have found transfer to the real world to be as good as, or better than, map study (Witmer, et al., 1995, 1996; Singer, Allen, et al., 1998).

Magee (1997) found transfer of ship handling skills from their Maritime Surface/Subsurface Virtual Reality Simulator (MARS VRS) to the real world. The system represented the bridge of a normal training ship and allowed instruction in low-level ship maneuvers. These maneuvers required following command and communication protocols, and giving orders which would maintain ship position in a formation under different weather and sea conditions. When VE-trained students were compared with ship-trained students on the trained maneuvers, during a delayed transfer test on-board training ships, the VE-trained students performed better. When more difficult, untrained maneuvers were introduced during shipboard training, no differences were found between the VE-trained and ship-trained students.

The major findings for spatial learning in this research support simulation conclusions drawn in previous research (e.g., Goldin & Thorndyke, 1982), and those found in reviews of simulation research (e.g., Hays & Singer, 1988). Since humans learn from interaction in a task environment, the necessary interaction should be supported. In spatial learning, this means that learners should be able to select their own routes, and experience the environment naturally by turning the head and looking around. Further, since humans learn by directly perceiving the environment, that environment should be adequately represented in VE. In other words, the fidelity of the representation should provide sufficient cues to the layout for learning. In learning cognitive skills like ship-handling the same rationale holds. The system must present sufficient perceptual cues for integration with cognitive rules and provide opportunities for practice and feedback.

Instructional Features in VE

Background

To date, we have conducted only one experiment addressing instructional strategies in VE-based training, and have not addressed instructional features at all. Instructional or training strategies vary the sequence or structure of the learning experience. Instructional features enable or support training strategies employed with training equipment. Instructional features are typically of two types. Instructional supports or aids serve to ease and improve the management of instruction, supporting or enabling the application of different strategies (e.g., after action review). Instructional adaptations enable changes in the presentation of stimuli used to initiate, guide, maintain, and/or provide feedback about the to-be-learned activity (Boldovici, 1992). These adaptations can take three basic forms; changing the stimuli through augmenting the normal stimulus or cue in the task environment, enhancing the normal cues by reducing the "noise" of conflicting stimuli, or by supplementing or adding an adjunct cue that calls attention to the normal stimulus.

Instructional support features have a long history of application in simulation systems (Hays & Singer, 1988). Conceptually they are sometimes difficult to classify and usually more difficult to investigate. This is because features are not independent and are almost always used in combinations in the support of some training strategy (Semple, et al., 1981). Features can be used to prepare, support delivery, and evaluate the success of instructional strategies. Some examples of instructional strategy support capabilities are automatic demonstrations; record and replay; after action reviews; automated cueing, scoring, and coaching; environment control; and control of problem insertions. Obviously, all of these functions can be incorporated in VE-based systems, depending on the use and goals of the system.

One of the greatest strengths of VE-based systems is the fact that all stimuli that are presented to the user are under complete control. As defined above, Instructional adaptations can be accomplished through changes in one or more of the stimulus dimensions (brightness, frequency, etc.) that are used to initiate, guide, or provide performance feedback information about the target activity (Boldovici, 1992). Instructional adaptations can also take the form of supplemental or adjunct stimuli, that is, additional stimuli or cues that call attention to normal stimuli used in the activity. For example, a circle of light may flash around an important cue in order for the student to better notice the cue during task performance. Finally, normally occurring non-task-relevant stimuli that might divert attention, or mask important task stimuli can be reduced in one or more of the stimulus dimensions (brightness, frequency, color, decibel level, etc.) in order to enhance the salience of the important task cues. Obviously the implementation of these functions is strongly related to the instructional strategy in terms of whether, when, and how much to use the adaptations, and the control of specific cues then requires instructional support functions.

Instructional Strategies in Spatial Learning

Although there are some widely accepted theoretical descriptions of how spatial knowledge might be acquired naturalistically, little research addresses the best strategies for training such knowledge efficiently. There is even less research exploring the most effective instructional strategies for training spatial knowledge in a VE. Bailey and Witmer (1994) assessed the effects of varying the amount of interactive exposure to the VE on the acquisition of route and configuration knowledge. Sixty-four participants rehearsed a complex route through a virtual building using an instructional strategy based either on finding and following successive landmarks (landmark-based strategy) or following right/left style directions (directions-based strategy). In addition, participants' FOV was either linked solely to body orientation (uncoupled) or controlled by both body orientation and head movements (coupled). Combining two levels of instructional strategy with two levels of head coupling created four levels of interactive exposure. The results indicated that the landmark-based training strategy produced more accurate route knowledge than the directions-based instructional strategy. Head coupling had no effects on route learning, but interacted with instructional strategy to influence one measure of configuration learning. Participants who used the landmark-based strategy performed better when head tracking was uncoupled. The head-coupled condition may have produced more disorientation and simulator sickness (because head and body orientation can vary independently) leading to poorer configuration learning in that condition. Support for this hypothesis comes from the fact that eight of the eleven participants who dropped out of the experiment due to simulator sickness were in the head-coupled condition.

In other research efforts, Darken and Sibert (1993; 1996) investigated the effects of various navigational aids in efficiently navigating a large virtual space. Aids included color coded grids that segmented the VE into distinct segments, virtual bread crumbs that marked areas already searched, a virtual sun for orientation, spatial audio that in effect enlarged the target object, and the provision of map views and flying views. They showed that using these aids improved performance on wayfinding, search, and configuration learning tasks. Although Darken and Sibert (1993, 1996) referred to these VE augmentations and supplements as navigational aids or tools, they are essentially instructional adaptations intended to improve the acquisition of spatial knowledge in VEs.

IV. Presence, Task Involvement, and Immersion

The sense of presence reported by users of VEs is what sets them apart from other simulated training environments. *Presence* is defined as the subjective experience of being in one place or environment (the computer-generated environment), even when one is physically situated in another (the actual physical locale). It is hypothesized that presence can make training in simulated environments seem more realistic and increase user involvement in the tasks and situations presented in the VE. The increased realism and involvement should improve task performance in the VE and promote transfer of VE-acquired skills to the real world. This definition does not explain the conditions under which presence may or may not occur, nor does it identify those factors that are likely to increase or decrease the amount of presence experienced in a VE. For dismounted soldier training, the most important issues are whether increased presence produces increased learning and transfer, and if so, how presence can be increased.

A Theoretical Framework for Presence

As discussed in Witmer & Singer (1998), presence depends on attention shifts from the physical environment to the VE, but does not require the total displacement of one's attention. Individuals experiencing a VE can attend to aspects of the VE and events in their physical environment concurrently. How sharply users focus their attention on the VE partially determines the extent to which they will become involved in that environment. We also contend that presence varies across a range of values that partly depends on the allocation of attentional resources, which determines the level of *involvement.* Presence also depends on the degree to which the VE includes *immersive* factors.

In general, as users focus more attention on the VE stimuli, they become more involved in the VE experience, which leads to increased presence in the VE. The amount of involvement will vary according to how well the activities and events attract and hold the observer's attention, and the extent to which the users are preoccupied with other problems or become focused on activities occurring outside of the VE. If the VE user is ill or the HMD is uncomfortable, involvement in the VE will likely be diminished accordingly.

A VE that produces a greater sense of immersion will produce higher levels of presence. Factors that affect immersion include isolation from the physical environment, perception of self-inclusion in the VE, natural modes of interaction and control, and the perception of self-movement. Aspects of VE that effectively isolate users from their physical environment (e.g., an HMD), depriving them of sensations provided by that environment, will increase the degree to which they feel immersed in the VE. If users perceive that they are outside of the simulated environment looking in (e.g., while viewing the environment via a CRT display), the immersive aspect is lost. To the extent that users find interaction with and control of a VE awkward, immersion in that VE is reduced. Perceiving oneself as moving inside a simulated environment or directly interacting with other entities in that environment will also increase one's sense of being immersed.

Factors Affecting Presence

Theoretical work by Sheridan (1992) and Held and Durlach (1992) suggested factors thought to underlie the concept of presence. In identifying factors and in developing items for measuring presence, we drew heavily on their work. Descriptions of the factors are included in Witmer & Singer (1994) and in Witmer and Singer (1998). While it is reasonable to assume that these factors may be associated with presence, considerable empirical work is necessary before we can confidently conclude that they do indeed affect presence.

Measuring Presence

Presence Questionnaires

We used those factors and others as the basis for the development of a Presence Questionnaire (PQ). In addition we developed an Immersive Tendencies Questionnaire (ITQ) to measure differences in the tendencies of individuals to experience presence. As part of our research program we used the questionnaires to evaluate relationships between reported presence and other research variables. Both questionnaires rely exclusively on self-report information. The PQ and ITQ use a 7-point scale format that is based on the semantic differential principle (Dyer, Matthews, Stulac, Wright, and Yudowitch, 1976).

Scale Reliability and Validity

Any measure of presence must be shown to be both reliable and valid. A measurement scale is *reliable* to the extent that individual differences in scale scores are attributable to true differences in the characteristics under consideration rather than resulting from errors due to random fluctuations in individuals or in testing conditions (Anastasi, 1968). A *reliable* scale consistently yields replicable scores. Reliabilities (Cronbach's Alpha) of different versions of the PQ have varied from .81 to .88, while those of the ITQ have varied from .75 to .81 (Witmer & Singer, 1998).

A scale is *valid* to the extent that it measures precisely what it purports to measure and measures it well. Items should cover a representative sample of the behavioral domain in order for a scale to have high content validity. PQ items were based largely on the factors derived from a review of the presence literature. Therefore their content reflects the conceptualization of presence as described in that literature. If the PQ is a valid measure of the presence construct, then PQ scores should be associated in a predictable manner with other variables or constructs that in theory are related to presence. A valid measure of presence should be associated in predictable ways with other variables and constructs.

Simulator sickness. In five of twelve experiments we found significant negative correlations between presence as measured by the PQ and SSQ scores. Nonsignificant negative correlations were found in six of the remaining seven experiments. Generally speaking, the significant SSQ-PQ correlations were found in experiments with higher SSQ scores. Based on these results, we argue that symptoms associated with moderate or severe levels of simulator

sickness (e.g., nausea, disorientation) draw attention away from the VE and focus that attention inward, decreasing involvement in the VE, thereby reducing the sense of presence.

Task performance. In only three of twelve experiments was PQ shown to be significantly related to measures of task performance in a VE. Statistically significant positive correlations between PQ scores and performance on simple psychomotor tasks performed in a VE were found in one experiment (Witmer & Singer, 1994), but not in a subsequent experiment involving the same tasks (Singer et al., 1995). Significant correlations were also found between PQ and performance on tests of spatial knowledge in two experiments (Bailey & Witmer, 1994; Singer, Allen, McDonald, & Fober, 1998), but not in two similar experiments (Witmer et al., 1996; Singer, Allen, McDonald, & Gildea, 1997).

The lack of a strong relationship between presence and performance found in our research does not preclude the existence of a strong association between presence and performance. Note also that these correlations between presence and performance do not necessarily indicate that high presence causes good performance, only that there is a statistical relationship. Good performance may cause high presence, or both may be produced by other factors. Controlled experiments are required in which a single subject group experiences several VEs that differ on how much presence they invoke. The average PQ score for the VEs would be the independent variable, and performance in those VEs or the amount of training transfer to the real world would be the dependent measure. The true relationship between presence and performance cannot be known until such experiments have been completed.

Natural modes of interaction. In one experiment (Bailey & Witmer, 1994), one group of participants could change their viewpoint in the VE by either turning their heads or by moving a joystick. For the other group, only the joystick controlled their viewpoint. Contrary to expectation, there were no significant group differences in presence as measured by the PQ. Similarly Lampton, Gildea, McDonald, & Kolasinski (1996) found no significant differences in the presence reported for head-tracked versus fixed view displays. In another experiment, Singer et al. (1995) found that different display views (stereoscopic versus monoscopic) interacted with head-tracking to change the perception of presence. Witmer & Kline (1998) found that more natural modes of moving through a VE such as walking on a treadmill produced significantly higher levels of presence than when movement was accomplished via a joystick. In contrast Singer et al. (1997) did not find significant differences in presence as a function of the method of locomotion used.

Relation to ITQ. If high ITQ scores reflect a greater tendency to become involved or immersed, then individuals who score high on the ITQ should report more presence on the PQ when exposed to a particular VE. Taken individually, only three of the twelve experiments resulted in a significant correlation between ITQ and PQ scores. The data across our first four experiments, which used version 2.0 of the PQ and ITQ, produced a significant positive correlation between ITQ and PQ scores (r=.24, p < .01). Similarly, combining the data from five of the remaining experiments, which used version 2.1 of the PQ and ITQ, produced a significant positive correlation (r=.15, p < .05 between ITQ and PQ. Experiments 4.1a and 4.1b were

excluded from this latter analysis because the range of PQ scores reported in these experiments was restricted to relatively low scores.

Conclusions regarding PQ construct validity. To date we have performed no research that directly compares the PQ scores of a single group of participants across vastly different VEs, nor have we investigated presence as a function of extended practice with a particular VE configuration. Nevertheless, our results to date suggest that presence, as measured by the PQ, is a valid construct. Supporting evidence includes a few significant positive correlations and no significant negative correlations between presence and VE task performance. Variables that might be expected to be significantly related to presence (e.g., ITQ scores, SSQ scores) were related, while variables that might be expected to be unrelated to presence (Spatial Ability Test scores, number of collisions in VE) did not correlate significantly with presence. The results of the cluster analysis also suggest a valid scale consisting of clearly differentiable subscales. However, we must emphasize that these findings are preliminary pending further testing and analyses.

The Role of Presence in Learning and Performance

While results relating measures of presence in VE to learning and performance in the VE and in the real world have been mixed (Witmer & Singer, 1994; Bailey & Witmer, 1994), many of the factors that appear to affect presence are known to enhance learning and performance. It is well established that meaningfulness and coherence of a stimulus set promotes learning (Underwood & Schulz, 1960). Gibson (1969) has suggested that selectively focusing one's attention on certain features of the environment to the exclusion of other features can be taken as evidence that perceptual learning has occurred. Bandura (1971) also reserves a prominent role for selective attention in his social learning theory. Learning is aided by requiring responses that are natural for the learner in a given situation (Seligman, 1970). As mentioned previously, interacting with the environment in a natural manner should increase immersion and hence presence. Factors believed to increase immersion, and hence presence, such as minimizing outside distractions and increasing active participation through perceived control over events in the environment, may also enhance learning and performance. Slater, et al. (1996) showed that participants who viewed tri-dimensional chess moves using an immersive display were more accurate in reproducing the moves in the real world than were participants who viewed the chess moves on a television screen. Because many of the factors involved in learning and performance also increase presence, it would be very surprising indeed if positive relationships between presence and performance were not found.

Clearly, presence is a multifaceted concept. It is not simply a matter of how involved an individual becomes in a situation or environment, though involvement is an essential component. Involvement, in turn, is affected by individual tendencies and selective attention. Our research with the PQ indicates that control, which affects immersion, is essential for a strong sense of presence. Other factors which are important include interface quality, sensory factors, naturalness of the interactions with VE, and simulator sickness. When interface quality is low or senses are not stimulated adequately, presence diminishes. Our research also suggests that moderate to severe levels of simulator sickness may reduce presence by distracting participants. Low levels of

sickness have no measurable effects on presence. The naturalness of the interactions with the VE and how closely these interactions mimic real world experiences affect how much presence is reported. For example walking on a treadmill to move through a VE produced higher levels of presence than moving through the VE via joystick manipulation (Witmer & Kline, 1998). Although we do not yet know precisely how the various contributing factors combine and interact to affect presence, we have begun to determine how they combine by identifying PQ and ITQ subscales. We maintain that both immersion and involvement are necessary for experiencing presence and that they interact to determine how much presence is reported. We do <u>not</u> claim to have identified all of the factors that affect presence, nor do we fully understand the presence construct, but we believe we have made considerable progress. We are continuing to refine the ITQ and the PQ as we learn more about the concept.

V. Simulator Sickness in Virtual Environments

In 1995, the committee on Virtual Reality Research and Design of the National Research Council outlined a series of recommendations they believed were crucial to the development of the synthetic environment field (Durlach and Mavor, 1995). One of the factors receiving the most emphasis was the issue of user comfort. They emphasized that if the comfort of VE systems, in particular HMDs, can not be radically improved then the practical usage of VE systems will be limited to emergency situations or to very short time periods.

In 1992, to support the planning and design of our research program and facilities, the senior author toured several VE research sites. From first hand experience and anecdotal reports, it became apparent that there was a potential problem of simulator sickness in VEs, and that there was little, if any, empirical data which would indicate the actual extent of the problem. Thus one of our early goals was to quantify the frequency and severity of simulator sickness in VEs. We have measured simulator sickness in each of our experiments. However, none of our experiments was specifically designed to investigate simulator sickness.

Simulator sickness refers to unwanted side effects and aftereffects that may result from using simulators such as flight or driver training simulators. Symptoms include nausea, dizziness, and headache or eyestrain. All discomfort resulting from use of a simulator should not necessarily be interpreted as simulator sickness (Kennedy et al., 1987). Simulator sickness refers to sickness or discomfort resulting from performing a task in a simulator for which performance of the same task in the real world does not produce similar sickness or discomfort.

Simulator sickness is a concern because of possible unwanted effects on the well being of the user and negative transfer of training. Simulator sickness may degrade training effectiveness despite the absence of severe symptoms such as vomiting. Discomfort in the simulator may distract the trainee. Simulator sickness may lead to negative transfer of training in that the trainees may adopt behaviors that mitigate sickness in the simulator, but will be detrimental if transferred to the actual task. Aftereffects involving the sense of balance, such as postural disequilibrium (ataxia) or visual flashbacks, could possibly impair the trainees' ability to drive safely or perform skilled motor tasks after leaving the simulator. The training value of a simulator is also reduced if simulator sickness forces a decrease in the frequency or duration of use of the simulator.

Basis of simulator sickness

Simulator sickness is thought to result, at least in part, because simulated movement results in a conflict between the human body's mechanical systems and visual systems for sensing movement. Treisman (1977) proposed that a change or conflict in the relationships between the senses might be interpreted by the body as an indication that toxins (poison) have been ingested. Therefore, nausea reaching the stage of vomiting would have survival value by removing the toxins. According to this explanation, simulator sickness is an unfortunate result of the inappropriate activation of this nausea response.

In the real world, our sense of balance is maintained by a complex interaction of several parts of the nervous system including: the eyes, which monitor where the body is in space (for example rightside up or upside down) and also directions of motion; the inner ears (also called the labyrinth), which monitor the directions of motion; the skin pressure receptors such as in the feet and seat, which indicate what part of the body is down and touching the ground; the muscle and joint sensory receptors, which indicate what parts of the body are moving; and the central nervous system (the brain and spinal cord), which processes information from the four other systems to make some coordinated sense out of it all. The symptoms of motion sickness and dizziness appear when the central nervous system receives conflicting messages from the other systems (American Academy of Otolaryngology, 1995).

In classic motion sickness, such as sea sickness in which the sufferer is inside the cabin of a heaving ship, the visual system does not register motion while the other systems do register that motion. Of more relevance to the consideration of simulator sickness is the situation in which the visual system indicates motion, but the other systems do not. Very wide FOV movie projection systems, such as Cinerama or IMAX, can produce compelling illusions of self-motion, called vection. Audience members have the visual sensation of riding in a roller coaster or helicopter but other sensory systems indicate that they are seated in a theater. Informal observations of audiences indicate that most viewers enjoy the experience, some experience more than mild nausea or dizziness, and a few viewers can not tolerate the experience and leave the theatre. Two aspects of this are highly relevant to simulator sickness. One is that visual displays alone, with no actual motion of the viewer, can produce symptoms of motion sickness. Hettinger and Riccio (1992) refer to this as visually induced motion sickness (VIMS). The second point is that there is a great deal of variance across individuals in susceptibility to VIMS.

In addition to VIMS, head and limb tracking systems may contribute to simulator sickness. Biocca (1992) discussed three ways in which imperfections in position tracking in HMDs can contribute to simulator sickness: lags in updating, especially when users move their head and the view of the VE drags noticeably behind their movement; jitters or oscillations; and conflicts between the visual display and the felt position. DiZio and Lackner (1992) pointed out that, independent of the visual display, the weight of an HMD changes the center of balance and inertia of head movements in ways that may produce symptoms of motion sickness.

As for the visual aspects of simulator sickness, Moffit (1997) noted that computer displays have routinely been associated with symptoms of visual discomfort including eyestrain, headaches, and blurred vision. He pointed out that HMDs have both complex eyepiece optics and electronic imagery and therefore are especially subject to multitudinous design or alignment problems that may lead to visual discomfort. Moffit concluded that HMD design, including adequate adjustability and an objective alignment method, could mitigate many of the problems of visual discomfort.

Kolasinski (1995) identified dozens of factors that previous research has indicated are involved in simulator sickness. Simulator sickness has been studied as a function of the

characteristics of the simulator, user, and tasks being simulated. Characteristics of the simulator include factors such as display image update rate and flicker. Characteristics of the user or trainee include age, gender, and ethnicity. Characteristics of the task being simulated include consideration of motion patterns such as rapid turns or slalom courses.

Table 5 (from Kolasinski, 1995) lists some of the factors that previous research has indicated affect simulator sickness. Note that the characteristics of the simulator itself are only part of the simulator sickness picture. The characteristics of the tasks being simulated and the characteristics of the trainees are also critical determinants of simulator sickness.

Table 4

Potential Factors Associated with Simulator Sickness in Virtual Environments

Individual	Simulator	Task
• age	• binocular viewing	• altitude above terrain
• concentration level	• calibration	• degree of control
• ethnicity	• color	• duration
• experience with real-world task	• contrast	• global visual flow
• experience with simulator (adaptation)	• field of view	• head movements
• flicker fusion frequency threshold	• flicker	• luminance level
• gender	• inter-pupillary distance	• method of movement
• illness and personal characteristics	• motion platform	• rate of linear or rotational acceleration
• mental rotation ability	• phosphor lag	• self-movement speed
• perceptual style	• position-tracking error	• sitting versus standing
• postural stability	• refresh rate	• type of application
	• resolution	• unusual maneuvers
	• scene content	• vection
	• time lag (transport delay)	
	• update rate (frame rate)	
	• viewing region	

Measuring simulator sickness

Questionnaires and symptom checklists are the usual means of measuring simulator sickness because there are many different symptoms of simulator sickness; measuring just one sign or symptom would not be sensitive (Kennedy & Fowlkes, 1992). A commonly used questionnaire to measure simulator sickness in flight simulators is SSQ (Kennedy, Lane, Berbaum, & Lilienthal, 1993).

The SSQ symptom list consists of 16 symptoms which are rated by the simulator user on a 4-point scale (0=none, 1=slight, 2=moderate, 3=severe). These ratings form the basis for three subscale scores - Nausea, Oculomotor Discomfort, Disorientation - as well as a Total Severity score. The symptoms making up the three factor scores are as follows: Nausea - general discomfort, increased salivation, sweating, nausea, difficulty concentrating, stomach awareness, and burping; Oculomotor - general discomfort, fatigue, headache, eyestrain, difficulty focusing, difficulty concentrating, and blurred vision; and Disorientation - difficulty focusing, nausea, fullness of head, blurred vision, dizzy (eyes open), dizzy (eyes closed), and vertigo. In addition to those 16 symptoms, there are 12 additional symptoms that were used in checklists that preceded the SSQ.

The Total Severity score uses all of the symptoms and reflects the overall extent of symptom severity. It is therefore the best index of whether or not a sickness problem exists. The SSQ subscale scores can provide diagnostic information as to the specific nature of the resulting sickness. Kennedy et al. (1993) have published baseline SSQ data obtained from Navy Flight simulators which can serve as a comparison for other systems.

In addition to the symptoms identified by the SSQ, loss of sense of balance, also called postural disequilibrium or ataxia, is another potential aftereffect of simulator exposure. Thomley, Kennedy, and Bittner (1986) suggested that ataxia may be caused by a disruption in balance and coordination resulting from the visual and vestibular adaptation to conflicting cues occurring during simulator exposure. Although sophisticated devices are being developed to measure ataxia, current research into simulator sickness often uses something similar to the "road sobriety test" administered by traffic officers. Wright (1995) has suggested that disorienting flashbacks may occur for some time after the exposure to the simulator.

Findings from our Research

Attrition

Simulator sickness symptoms have been sufficiently severe that 5.6% (29/517) of the participants have withdrawn from our experiments prior to completion. Rates for individual experiments varied from 0% to 25%. This is generally consistent with other findings. Other researchers have reported withdrawal rates of 10% (Garris-Reif and Franz, 1995) and 0% to 5 % (Regan, 1995). The experiments with high attrition rates tended to involve frequent self-motion and periods of constant exposure to VE (i.e., between breaks) greater than 10 minutes. The mean

SSQ Total Severity score for those participants who withdrew from our experiments was 79, while for those who did not withdraw it was 21.

Preexperimental Symptoms

Unfortunately, the symptoms of simulator sickness can also result from many other conditions or activities, such as colds, flu, consumption of alcohol, or even insufficient sleep. As a consequence, participants frequently report symptoms prior to the start of the experiment. Sixty percent of the preimmersion SSQs listed at least one symptom, and 9% listed a moderate or severe symptom. Fatigue and drowsiness were the most frequently reported symptoms on the preimmersion SSQs, being reported by 32% and 29.4%, respectively, of the participants.

The reporting of symptoms prior to immersion is not unique to college students. Lampton, Kraemer, Kolasinski, and Knerr (1995) report mean SSQ Total Severity scores for tank driver trainees of approximately 5.0 prior to training sessions. Regan and Price (1993) found that 54% of their sample of civilians, military personnel, and firefighters reported symptoms on their preimmersion checklist.

The frequent occurrence of preexisting symptoms raises a methodological issue in simulator sickness measurement. Measuring symptoms prior to immersion is necessary to determine the effect of the exposure on the level of postimmersion symptoms. Not only may preimmersion problems carry over to post immersion, they may increase susceptibility to simulator sickness. Alternatively, there is concern that seeing the list of symptoms beforehand may sensitize the participant to those symptoms or in some other way lead to exaggerated reporting of simulator sickness. Lampton, Kraemer, Kolasinski, and Knerr (1995) found no significant difference in postimmersion SSQ scores between groups of participants who were administered a preimmersion SSQ and those who were not.

Postexperimental Symptoms

Fifteen percent of the postimmersion SSQs (first 16 items) listed no symptoms, 52% listed only mild symptoms, 28% listed at least one moderate symptom, and 4.6% rated at least one symptom as severe. Eyestrain was by far the most frequent symptom (reported by 54%), followed by general discomfort (46%) and fatigue (45%). These three symptoms, along with headache and nausea, also showed the greatest increase (postimmersion - preimmersion) in the percentage of participants reporting them. Some level of nausea was reported on 22% of the post SSQs, and at 2.9%, nausea was rated "severe" more frequently than any other symptom. "Dizzy eyes open" and dizzy eyes closed" had very similar frequency ratings. Slightly less than 20% listed any dizziness, and less than 1% rated dizziness as severe.

Fourteen percent of the 338 SSQs that included the additional items listed in part 2 of the table indicated vomiting. Eight of these were from experiment 2.2, four from 5.1, and two from 1d. Ten of the fourteen withdrew from the experiment.

Distribution of Total Severity Scores

The pre- and postexperimental distribution of SSQ Total Severity scores is shown in Figure 3. As a result of experimental participation, Mean Total Severity increased from 8.09 to 24.84. Perhaps more importantly, the percentage of participants scoring above 20 increased from 11.3% to 42.2%. A total severity score of 20 is considered by Kennedy et al. (1992) to be one which requires retention of participants until symptoms have subsided.

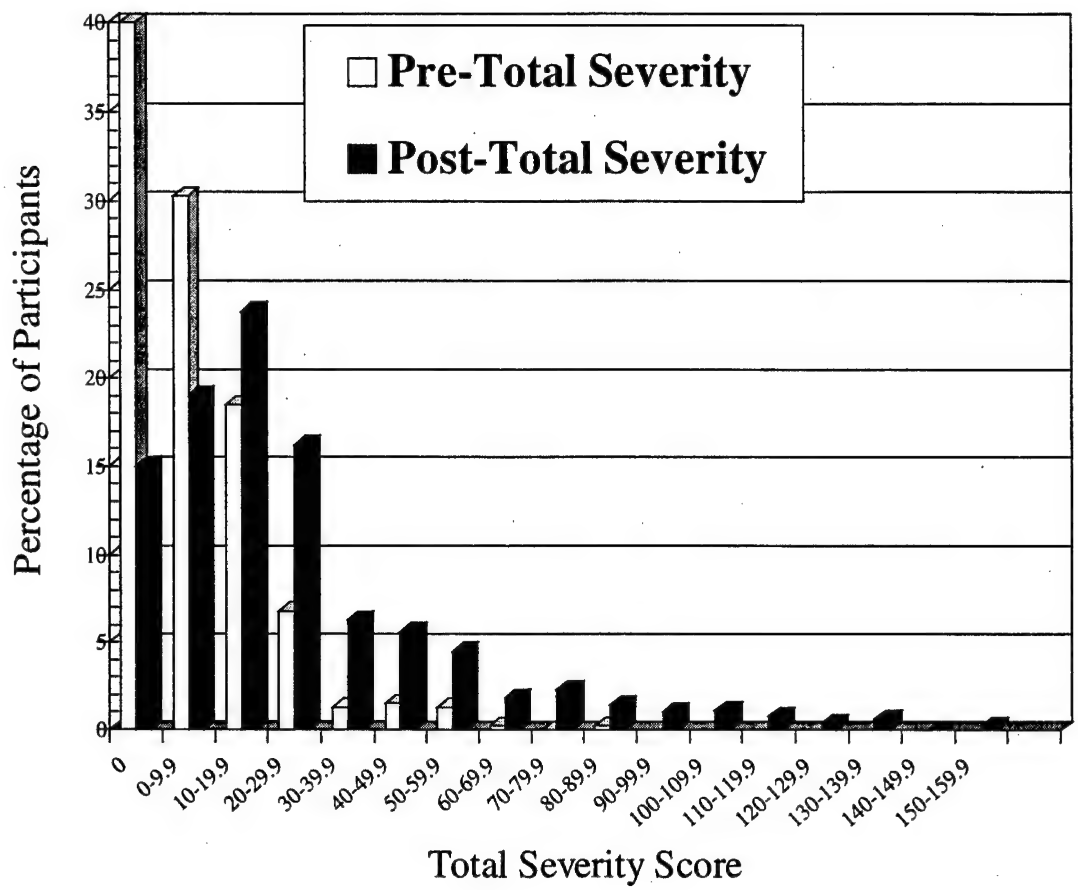

Figure 3. Distribution of pre- and post-SSQ Total Severity scores.

Total Severity Scores by Experiment

Measures of SSQ Total Severity Score central tendency for each experiment are shown in Table 6.The experiments with high average SSQ total severity scores tended to involve frequent self-motion, collisions, and periods of constant exposure to VE (i.e., between breaks) greater than 10 minutes.

Table 5

Mean SSQ Total Severity Score by Experiment

Experiment No.	N	Mean	S.D.
1.1 (day 1)	23	26.18	25.34
1.1 (day 2)	20	38.15	32.93
1.2	47	15.12	15.53
1.3	49	18.32	19.33
2.1	24	37.86	39.34
2.2	74	32.40	33.48
3.1	52	31.01	28.05
3.2	34	23.32	24.69
4.1a	24	13.09	13.69
4.1b	28	13.36	15.52
4.2	71	15.22	15.77
4.4	24	28.05	26.18
5.1	42	33.04	27.10

Ataxia

Many of our experiments have included an HMD with head tracking. We have attempted to use head tracking as a way to measure postural stability following the VE exposure. Participants stood as still as possible on one leg while wearing the HMD but with no visual display being presented. We have not found significant changes in the amount of head movement as a result of exposure to VE. Either ataxia is not occurring, or the system is not sensitive enough to pick up slight head movements. Grant and Magee (1997), in their science center experiment, assessed postural stability before and after the experiment using a force platform. They found no significant effects due to VE exposure.

Effects of Multiple Exposures

Of particular relevance to the use of VE for training applications, are the findings in previous research, mostly involving flight simulators, indicating that, all other things being equal, a trainee is most susceptible to simulator sickness during the first session with a simulator. For most trainees, simulator sickness declines during subsequent sessions. Only two of our experiments (1.1 and 1.3) have involved multisession exposures to VE. In experiment 1.1, symptoms were more severe in the second session than in the first, perhaps because it was of longer duration and the tasks involved more self motion (Lampton, et al, 1994). Experiment 1.3 data showed no discernable pattern across sessions. Data from the M1TDT training program and the DWN engineering experiments (Lockheed Martin, 1997) indicate that users reported fewer and less severe symptoms of simulator sickness with subsequent training sessions in the simulator. Regan (1995) also report decreasing symptoms over a series of four exposures.

Time Course of Simulator Sickness within and across Immersion Sessions

Most of the participants who have withdrawn from our experiments because of simulator sickness have done so within the first 5 minutes of immersion. For those who do not withdraw, the severity of symptoms seems to accumulate, at least up to a point, over time of immersion. Regan and Price (1993) found that reports of symptoms increased during each of four five-minute intervals across a 20-minute immersion session, then decreased during two five-minute recovery periods. Singer et al. (1998) measured simulator sickness before, midway through, after an average of 30 minutes of immersion, and after a 30-minute recovery period. They found that midpoint SSQ scores were significantly higher than the preimmersion scores, but the midpoint and immediate postimmersion scores did not differ. Post-recovery period scores did not differ from the preimmersion scores, indicating that recovery occurred during a 30-minute waiting period.

Effectiveness of Preventative Medications

Regan (1995) administered a nonprescription, antimotion sickness medication (hyoscine) to participants prior to immersion in VE. They found a reduction in nausea, headaches, and eyestrain. An unanswered question is whether medications can be used to reduce symptoms without also affecting the trainee's learning ability.

Predicting Individual Susceptibility

As part of the background questionnaire, participants rated themselves as having a history of motion sickness (yes or no) and, on a seven point scale, their susceptibility to motion sickness. Correlations of these items with postexperimental SSQ Total Severity were significant but small (.18 and .16, respectively). In contrast, Lampton et al. (1995) reported correlations of .50 and .43 for history and susceptibility, respectively, with SSQ Total Severity following first use of the M1TDT. Perhaps their correlations were higher because the motion base of the M1TDT produces cues that cause motion sickness, as well as simulator sickness. Our data did show a significant correlation between pre-and postexperimental Total Severity (r=.32) very similar to that found by Lampton et al. (r=.38). This may show, as some theorists believe, that those who are already symptomatic are more likely to develop simulator sickness, or it may indicate only that individuals differ consistently in their willingness to report symptoms.

Practices Adopted

During the course of our experimentation, we have adopted a number of practices to reduce simulator sickness among our subjects. While these have not been subject to formal experimentation, they are based on pilot testing and our own personal experience and reactions:

For the first exposure, use short sessions (10 - 15 minutes) interspersed with breaks (5-10 minutes). If the participant is standing, provide a stable surface for hand contact. The room should be cooler than normal, with fans to provide air movement. It may or may not reduce symptoms, but it makes them more tolerable. The fan cools the participant, provides the

sensation of "fresh air," and provides a means to give a directional cue to "real world" orientation without introduction of a distracting visual cue in the VE.

Simulator Sickness Measurement Issues

The SSQ total severity scale developed by Kennedy, Lane, et al. (1993) is a valuable tool for measuring simulator sickness and provides a common denominator for comparing simulator sickness across different experiments. The SSQ subscales were the result of a factor analysis of simulator sickness data from Navy flight simulators. The subscales were useful for interpreting the data on which they were developed. However, they may not be appropriate for use with different populations (soldiers and college students instead of Navy pilots) and different conditions (VE instead of flight simulators).

There are (at least) two potential problems in applying the SSQ subscale formulae for interpreting our simulator sickness data and VE data in general. The first involves the weightings of the subscales. The weights were selected to produce equivalent standard deviations of 15 for the subscales based on the Navy simulator sickness data. When applied to our data, they produced standard deviations of 34.5, 25.6, and 21.1, respectively, for the disorientation, nausea, and oculomotor subscales. The unweighted scores had standard deviations of 2.48, 2,69, and 2.79. Weighted means were 24.0, 18.7, and 22.4 , respectively, for the disorientation, nausea, and oculomotor scales, while the unweighted means were 1.72, 1.95, and 2.95. The weights therefore changed the relative order of the means and made direct comparisons among the subscales more, rather than less, difficult. The unweighted scores also lend themselves to quick interpretation (a score of 1 equals one mild symptom, 3 equals one severe symptom or three mild symptoms, etc.) which is not possible with the weighted scores.

The second potential problem involves the interrelationships among the subscales. Some symptoms (general discomfort, nausea, difficulty focusing, difficulty concentrating, and blurred vision) are included in two subscales. For example, the item "nausea" is used in computing both the nausea and disorientation subscale scores. This contributes to high correlations among the subscales (.65 to .76 in our data), and reduces the usefulness of the subscales as a diagnostic tool. Both of these problems suggest a need for new scoring procedures for the symptoms which result from the use of VE.

We have slightly modified the SSQ by replacing the "sweating" symptom item with two items, "warm sweating" and "cold sweating." This was done to prevent SSQ scores from being artificially inflated by the sweating that results from normal physical exertion. Unlike the Navy flight simulators in which trainees remain seated, simulators for dismounted combatants may involve walking or running. Therefore, it is desirable to differentiate warm sweating, which may occur from physical exertion or warm room temperature, from "cold sweating" which may result from nervousness or simulator sickness. If the VE system involves physical exertion (for example, if a treadmill is used for locomotion), then the SSQ should have two separate items for rating sweating "Warm" sweating (due to exertion) versus "cold sweating" resulting from discomfort or nervousness. Only "cold sweating" is actually used in calculating SSQ scores.

VI. Hardware and Software Issues

In the course of our program, a number of VEs have been designed and developed. The experiments described in this report have used three different image generators and a wide variety of different display and interface devices. As a result a number of problems have been encountered and solutions developed. This section describes some of the key lessons learned from that experience.

Hardware Performance Specifications Based on Ideal Conditions

While there is no suggestion that manufacturers deliberately misrepresent the capabilities of their equipment, they nevertheless do present its specifications and performance characteristics in the best possible light. Consequently, those data usually represent the best possible performance under ideal conditions, and not the performance that is to be expected under realistic conditions of use. We will illustrate with the following two examples.

Rinalducci, Mapes, Cinq-Mars, and Higgins (1996) measured the visual display characteristics of two HMDs from the same manufacturer. While the manufacturer indicated that the visual displays contained 360 X 240 pixels, the nature of the optics in both displays was such that less than two-thirds of those pixels were visible to the user. Also, in one of the HMDs, the horizontal FOV was ten degrees smaller than indicated.

During the DWN experiments (Lockheed Martin, 1997), soldiers using two VICS were asked to fire at virtual targets while in a prone position. This should have been possible according to manufacturer's specifications for the weapons tracking systems used. After 15 trials it was obvious that neither system could engage targets reliably from the prone position. In a system using acoustic tracking of weapon position aiming error was so great that in many instances the direction the weapon was pointing was reversed and the soldier ended up hitting himself rather than the target. While the system had previously worked acceptably in a carpeted laboratory, the high level of ambient noise at the experimental site, where multiple systems and personnel were located in one large room, substantially degraded performance. A system that used video tracking lost track of markers and mis-assigned weapon and soldier body part markers. The value of testing the performance of system components (hardware or software) early in the development process should be obvious.

Limitations of Current Tracking Technology

This section focuses on the limitations of current electromagnetic trackers, since that is the type of technology that we have used. However, electromagnetic trackers are not the only ones with limitations. As described above, visual and acoustic trackers are also subject to interference.

Magnetic position tracking suffers from two major drawbacks: latency and interference from metal in the vicinity of the tracker. Latency can be abated by configuring the tracker with its

own computer or processor. This allows the tracker to operate in 'streaming' mode, where position and orientation data is supplied constantly, and a dedicated processing unit handles only the task of receiving the data and preparing it for use by the principal simulation process. In this configuration, latency issues are not eliminated completely, but reduced to a manageable level.

Problems with environmental interference are more difficult to overcome. When a receiver is close to a metallic object, the receiver sends back faulty position and orientation data to the tracker collection unit. Depending on the size of the metal object and its position relative to the transmitter and the receiver, position information may be off as much as 12" along any axis, as well as orientation errors up to 10 degrees! Sensitivity of the trackers to metal in the environment is great enough that small groups of people with coins and keys in their pockets and purses are sufficient to interfere with gesture recognition.

A problem discovered using an electromagnetic tracking system for gesture recognition is that while the electromagnetic sensors may provide a continuous view of three-dimensional space, that space is not necessarily linear. This presents problems because if the field is "pinched" in a specific area, readings outside of this area would be normal, but readings inside this area would be exaggerated in some manner. This has a profound effect on recognition because as space "bends," the degree of accuracy degrades, creating recognition errors. Making all orientations relative did not solve the problem. For example, if the participant holds out their arms and a reading is made, then rotates their entire body by 90 degrees and a second reading is made, the *relative* orientations calculated will be different. In reality, the actual relative orientations have not changed, but it is impossible to determine this from the information provided by the sensor. This is because sensors closer to the transmitter provide more accurate position information than sensors farther away. The linearity of space also varies over time; during similar conditions on separate days, the linearity of space as perceived by the distance reading between two sensors has shifted by as much as 25% or as little as %0.2. Such shifts are due to invisible conditions such as fields generated by unshielded monitors.

While preparing for the body model experiment, the tracker readings were determined to be very close to the actual measurement values with one notable exception. The roll component varied plus or minus 4 degrees as the tracker was rotated 360 degrees on a level plane. This was very apparent to users of the HMD as they performed a 360-degree turn. The roll component of the HMD movement was turned off for this experiment, rather than perform a lengthy calibration table process.

The problem is compounded by the difficulty of forecasting what type of interference might be encountered in a particular environment. The only certain way of resolving the effects of tracker interference is to very carefully (and tediously) measure the tracker readings across a uniform space with known distances and orientations, such as a marked grid, and use the readings to tabulate a correction table. This table must be measured with granularity as fine as the rating of the tracker, otherwise an interpolated method must be used. If the transmitter is shifted even slightly in position or orientation, or if the position of metallic objects within the tracker environment changes, the correction table becomes useless.

Post-Processing of Performance Data is Recommended when Extensive Processing is Required

The underlying goal of software engineering for real-time immersive simulation is to achieve the highest frame rate possible. One method for maintaining a high frame rate is to offload non-critical central processing unit (CPU) tasks such as data processing. Rather than performing statistical analysis on data during the experiment, data is best written to a terse, un-annotated intermediate file during simulation. These terse data files are designed to be imported directly into statistical analysis packages for later analysis. The less explanatory textual information (the more terse the intermediate file) the better, as each superfluous character written to file during run-time lowers overall efficiency of the experiment's application process. By only writing terse data during simulation, flexibility in the statistical analysis is possible. Fundamental experimental data written in terse format can be analyzed in multiple ways, allowing experimenters to adjust their analysis according to experimental results, as opposed to 'hard-coded' analysis that occurs during run-time.

Special Interface Problems and Solutions

Creativity in the design of VE interfaces is required to develop the experimental settings required for research. IST has modified treadmills, developed a special "pod" for walking, and designed and built a number of safety harnesses and constraints. In addition to the pod, which has already been discussed, two small interface devices appear to be applicable to training as well as research settings.

IST student researcher Trent Tuggle developed the "Tugglestick" as a low-cost alternative to the $1200.00 Silicon Graphics, Inc. (SGI) Flybox joystick controller. A PC Joystick was wired directly to a Motorola 68HC11 Evaluator card. The tugglestick integrates the two buttons on a basic PC joystick into the controller as well as the joystick itself. The Motorola card was configured as a UART RS-232 serial device compatible with the serial connections available on the SGI RealityEngine architecture. An SGI-compatible device driver for the card was then written and integrated into the system. Total hardware cost was $80.00.

The "Appleratus" was designed to be a comfortable hand-held input device. It is made of wood, lightweight, and shaped like an apple; in a previous incarnation (before it was transformed into a high-tech input device) it was a craft store apple. Incorporated in and on the Appleratus are a four-wire modular phone jack, two push buttons, and an electromagnetic tracker. A standard modular phone cable connects the Appleratus to the Tugglestick. The two buttons recessed within the front surface of the Appleratus appear to the tugglestick controller to be joystick buttons built into the device. The Appleratus was designed to use the same circuitry and assembly code that the tugglestick uses. This allows the software driver for the tugglestick to also be used for the Appleratus. The electromagnetic sensor tracks the position of the Appleratus, enabling it to be used as a pointing device or to simulate a hand-held weapon.

For team training, a more sophisticated manual control device was adapted from a gaming system joystick. A button under the participant's thumb enables toggling through the set of equipment available during a mission scenario. A button under the index finger activates the

currently selected equipment. For example, a thumb press toggles from empty hand to pistol and then the index finger button fires the pistol. Preliminary testing indicates that participants quickly learn to use the manual control device.

Need for Multiprocessor Systems

As the experimental suite has reached a level of complexity capable of supporting a fully-immersive multi-player environment for team training experiments, it has become apparent that the computing power of single processor computers is insufficient to sustain multiple subjects in an immersive environment. In addition to the network overhead necessary to support a multi-player synthetic environment, serial input devices (such as magnetic position trackers and the Appleratus) play an important role. These tasks require many CPU cycles at a rate necessary to sustain an immersive experience. Consideration must also be given to the many application tasks that we require our hardware to perform in addition to fundamental image rendering tasks. These tasks include computationally expensive requirements such as intersection testing used for collision, terrain following, and avatar/world and avatar/avatar interactions. Other common tasks that require large amounts of CPU cycles are dynamic database loading, network Protocol Data Units (PDU) processing, data collection, spatial audio calculations and general sensor processing and updating. The Onyx RealityEngine can be configured with a variable number of processors, and load balances tasks efficiently by running network and asynchronous Iinput/Output (serial I/O) functions on a processor separate from the processors responsible for the primary simulation tasks.

The graphics capabilities of PCs are beginning to show promise. The MMX and Advanced Graphics Port (AGP) capabilities available with the Pentium II chipset allow geometry and texture data to migrate from the general bus to a bus dedicated just to graphics with a throughput potential of 512 Mbytes per second, or about 25% of the speed of the RealityEngine principle graphics bus. Despite the fact that the RealityEngine does not normally move textures along its bus, but stores them in special fast memory units within the raster managers, the speed of current PC AGP buses are beginning to approach the speed of RealityEngine hardware.

Where PCs fall short, however, is in their ability to take advantage of multiple processors. As mentioned above, achieving adequate simulation frame rates for a fully immersive synthetic environment application depends on more than geometry throughput. PC architectures today are currently available with two processors, and soon architectures with more processors will be on the market. In order to exploit these processors to their fullest, an operating system for PC's and real-time simulation Application Programming Interfaces (APIs) is needed that maximizes the potential of multiprocessor architectures.

Windows NT is not a true multiprocessing/multitasking operating system. It is multi-threaded in its latest incarnation. Linux, a freeware operating system for PCs that is very similar to Unix, is the only PC operating system available today that is truly multiprocessor capable. In addition, PCs lack the multiprocessor support that the high-end graphics workstations offer developers embedded within their graphics APIs to load balance simulation tasks (such as application tasks, draw list preparation, culling scene elements to the view frustum, and

intersection/collision traversal) efficiently. This is why the use of high-end PCs is still insufficient for immersive environment applications that require more than rudimentary fidelity or utilize multiple complex application tasks. As long as PCs remain directed at the home/office/entertainment markets, and because multiprocessor rendering platforms are critical to the continuation of ARI dismounted infantry training experiments, development will likely continue to be done on graphics workstations for at least a few more years.

Use of Proprietary Software

Software for all experiments prior to 3.1 were performed on either PC-based systems or on an SGI Crimson RE platform using a commercially available development API. When the software was migrated to higher-end multiple-processor systems (the Onyx, a four-processor system), these APIs did not provide for multiprocessor support. They were capable of running on multiprocessor systems, but only took advantage of, and used, a single processor. Since the demands of the experiments were increasing and requiring more processing time, taking advantage of multiple processes was imperative. We moved to a then-recent release of the Performer API from SGI because it was finely tuned and highly optimized to the SGI architecture. Development is done with an object oriented programming language. This allows low level, fundamental sections of code that are frequently used, such as device drivers for the tugglestick, treadmill, HMDs and the like, to be coded as atomic units that do not use API calls. This provides more flexibility when moving software to different computing platforms. Licensing costs also played a role in the decision to move from a commercially purchased API to developing an in-house software library based on Performer.

Human Figure Models and Communication via DIS Protocols

Fully articulated real-time human figures are not natively or explicitly supported by the latest release of DIS. For the team training experiment, developers used articulated parts PDUs attached to an Entity Service PDU for transmitting the angles of body components. The team training body model has 47 moving parts, but not all of these are used for the experiment. For example, both arms of the model are fully articulated, but sensors are only available for one arm to be tracked. Rather than send an articulated part PDU for a body component that has not changed position since the last transmission, the parts that have changed since the last transmission are recorded in a bitmask that the receiving clients decode. This allows faster processing and reduced bandwidth requirements.

Use of PCs for Sound Recording and Playback

During team training experiments, communication between the subjects is essential, and plays an important role in mission success. All communication that occurs among subject 1, subject 2 and the experimenter are captured as .wav files on a PC running Windows 95. The audio capture program uses a volume threshold setting to determine whether or not to capture sound data. This allows shorter, multiple sound files to be generated, rather than one long file that is very expensive in terms of disk storage space. These sets of files pertaining to a mission can be played back in sync with the visual playback for after-action critique purposes.

VII. Recommendations

Introduction

Our recommendations for the use of VE for dismounted soldier training should be viewed in the overall context of what VE will be used to train dismounted soldiers to do, and how it will be used. VE training systems are likely to remain expensive to acquire, operate, and support for the foreseeable future. For that reason their use will be limited. Few units will have access to a system that can support an entire unit (e.g., platoon) simultaneously for training purposes. In the training domain, we see the following uses for VE: preparation for specific missions (mission rehearsal); training small unit leaders; and training in the use of digital systems.

- Full units using VE simultaneously will likely be limited to units preparing for critical or dangerous missions (e.g., hostage rescue, peacekeeping). Soldiers will be able to use computer-generated models of the actual terrain and structures where the mission will be conducted to learn to recognize landmarks and routes, identify danger areas, and determine lines of sight and fields of fire. Leaders will be able to refine their plans and better prepare for different enemy situations.

- Small unit leaders will require additional and better training as their tasks become more difficult. VE, using simulated friendly and enemy soldiers, can be used to train leaders in the cognitive skills required for MOUT, night operations, and peacekeeping without requiring the support of their and other units. This will be most beneficial in the early stages of training.

- Digital systems for the dismounted soldier require training both in their operation and the tactics and techniques for using them in combat. VE can be used to train the cognitive skills for their effective use. This will also be beneficial in the early stages of training.

The recommendations in this section are derived from the research presented in the previous sections of this report. In some cases a recommendation follows directly from the results of a single experiment. More frequently, the recommendations are derived from the combined results of multiple experiments. The recommendations have been organized into the following categories: interface design; improving distance estimation; tasks suitable for training in VE; improving training and training transfer; presence; reducing simulator sickness; and recommendations for researchers.

Interface Design

Trainees should always receive training and practice in the use of the interface they are to use before they begin training on the target task or tasks. The goal of this interface training should be to insure that all trainees meet defined minimum standards for the use of the interface.

Visual Displays

The use of a bi-ocular display (presenting the same image to each eye), rather than a true stereoscopic display, will improve the scene update rate, or conversely, permit the presentation of more visual detail, without adversely affecting distance perception if the distances involved are relatively long. It may also reduce eyestrain.

Relatively inexpensive HMDs provide poor visual resolution relative to human visual capabilities. An HMD with VGA resolution (640x480) and a 46-degree horizontal FOV provides a resolution of about 4.3 arc minutes per pixel (or triad). The normal resolution of the human eye is about one arc minute. These devices are not appropriate where fine visual discriminations (such as identifying targets at long distances) are required.

Low-end HMDs may also not represent color accurately. Successful task performance or training should not require fine color discriminations.

The actual FOV of the display device and the geometric field of view (GFOV) of the software should be the same. Otherwise errors in distances estimation may result. If the GFOV is set to a value other than the actual FOV, trainees should be informed of that fact.

The use of a self-image may or may not improve individual task performance. However, it is required for training team tasks where visual contact among team members is required.

Simulating Locomotion

The process of moving through the VE should require minimal attention on the part of the user. If the task requires use of the hands to perform other tasks simultaneously or nearly simultaneously with movement, then some form of foot-based locomotion control will be required. If the simultaneous use of the hands to perform other tasks is not required, a well-practiced method, which is not like the real world (such as a joystick), may be preferable to a novel method that is more similar to actual walking. A well-practiced method of locomotion, even if it is not similar to walking in the real world, may require less of the trainee's attention than would an unfamiliar but more realistic walking technique. Use of the well-practiced method would permit the trainee to pay more attention to acquiring the skills to be trained, and less to the process of locomotion. Which technique is better in a particular situation will depend on the trainee population's skills in the use of the well-practiced method, the walking technique used, and the tasks to be trained.

Trainees seem to learn to move through VE using any of a variety of well-designed interface devices. College students, junior Army enlisted personnel, and company grade officers are particularly proficient in the use of a joystick, perhaps because of experience with video games and computer games.

If trainees are to assume different positions (i.e., standing, prone) and the interface is such that the trainees' positions in the real world do not always correspond to their positions in the virtual

world, the trainee should have an easy way to determine his position in the virtual world (such as an Icon).

"Teleporting," to locations not in the line of sight eliminates cues about spatial configuration and distance traveled. If one training objective is to impart spatial knowledge, teleporting should be avoided. If egocentric orientation is to be maintained in the VE, teleporting may also lead to self-location delays or errors.

Manipulation

Current tracking technology is not sufficiently accurate to simulate the precision use of small arms, or to track rapid or precise movements.

Many forms of interaction in VE can be accomplished by pointing or "grabbing" and "dragging" with relatively natural positioning motions based on whole hand or end of limb position (with a single sensor on the hand).

Use of workarounds

If the workaround (e.g., for locomotion or manipulation) is adapted from a well-learned skill (for example, walking in place for walking), then there is little risk that it will interfere with subsequent task performance in the real world.

Improving Distance Estimation

For tasks involving/requiring accurate estimation of distance, use the widest FOV display device available. Both horizontal and vertical FOV may be important. If a wide FOV display is not available, consider using compensatory cues or some similar method for calibrating distance judgments in the VE.

If accurate distance estimation in VE is required, either provide training on this task before training the target task, or provide compensatory or augmented cues during the performance of the task itself.

Head-coupling and binocular displays both improve distance estimation, but only at relatively short distances (ten feet or less).

Tasks Suitable for Training in VE

Individuals can learn to find their way through real world spaces by rehearsing in virtual representations of those spaces. Currently it may be more efficient to rehearse in the real environment when possible, but if the real environment is not available for whatever reason, then VEs are recommended as the best substitute for learning about complex spaces.

VEs can also be used to train configuration knowledge as well as route knowledge effectively, but additional work is needed to determine how VEs can be used to their best advantage in training this type of spatial knowledge.

VEs are not currently suitable for training psychomotor tasks, particularly those that involve precise or rapid motor activity. Systems that use HMDs are also not well suited to tasks that require rapid head movements.

The use of VEs for sustainment training of psychomotor tasks that require precise multi-modal feedback requires periodic performance evaluations to insure that the absence of some forms of feedback is not causing performance to deteriorate.

VEs are not currently suitable for training tasks that require accurate use of individual weapons.

Improving Training and Training Transfer

Disorientation in VEs is a major impediment to training spatial skills efficiently in those environments. While wider FOV VE displays and more natural VE movement paradigms would alleviate some of the disorientation, other alternatives for reducing VE disorientation should be investigated.

Landmark-based training strategies and navigational aids have proven useful in training spatial knowledge in VE and should be utilized to their maximum advantage in VE spatial training.

For geo-specific training, visuals that are not known to be accurate should be obviously generic, so trainees do not learn incorrect visual cues.

Presence

Presence is weakly but positively related to performance. To maximize performance in a VE, first design the VE and tasks in conformance with good training principles. Then increase the factors that contribute to experiencing more presence.

Moderate to severe simulator sickness decreases the amount of presence experienced and reported. Reduce or eliminate factors that produce moderate to severe levels of simulator sickness to increase presence.

Reducing Simulator Sickness

Simulator sickness is a real but manageable problem. Its reduction should be a major consideration in the design and development of a VE-based training system, and in determining the environment, activities, and duration of the trainee's first use of the VE.

During the development of the system, periodic "test drives" should be conducted from the perspective of detecting problems with simulator sickness. Those who should conduct these test

drives should be representative of the expected users. In particular, if the trainees will be naïve with regard to the training system, the testers should be naïve as well.

Simulator sickness should be a primary consideration in determining the environment, activities, and duration of the user/trainee's first immersion in the VE.

A trainee's initial immersion in the VE should:

Be brief (inoculate against the simulator sickness). For the first exposure, use short sessions (10 - 15 minutes) interspersed with breaks (5-10 minutes)

Be designed to prevent sickness producing activities such as rapidly slewing FOV, collisions, and viewing objects at very short distances. The activities should not require or encourage rapid head motions.

Avoid depicting movement not under the participant's control.

The first session should not include a lengthy period of "free play" in which the participants becomes familiar with the VE interface by doing whatever they want.

Participants should always be alerted to the possibility of simulator sickness, how to recognize the onset of symptoms, and what to do if they begin to become ill.

Periodically check the calibration of the system.

To check that appropriate levels of resolution and color presentation are being maintained provide a simple "test pattern" for the user to inspect. A naïve participant will not recognize that an HMD is not operating as expected.

Require the participant to speak frequently during train-up. Pallor (extreme paleness) of the participant may be an indicator that simulator sickness is developing. However, An HMD may occlude much of the participant's face. Requiring the participant to speak may provide change in voice cues that simulator sickness problems are developing.

The tremendous variability across individuals in susceptibility to simulator sickness has several implications: For the simulator sickness "test drives" use more than one tester. (Just because your VE system does not make you sick does not mean some of your participants will not have severe symptoms.) Even with systems for which most participants do not experience severe symptoms, some individuals may become nauseated to the point of vomiting. Airsickness bags or such should be readily accessible.

The room should be cooler than normal, with fans to provide air movement.

If the participant is standing, provide a stable surface for hand contact.

Be aware that the SSQ subscale scores are potentially misleading in that the weights may not be appropriate for your simulator sickness data distributions.

Avoid very fast optical flow. For example, if you simulate fast crawling, do not attempt to depict in detail what dirt looks like from 3 inches away.

Plan to retain trainees in the training area until symptoms subside.

<u>Recommendations for Researchers</u>

The most accurate unbiased distance estimates are obtained using the Non-visually Guided Locomotion (NVGL) procedure and are recommended for obtaining distance estimates in both VE and in the real world.

Presence requires both involvement and immersion. Therefore measures of presence must include items addressing both of these factors.

The ITQ predicts PQ scores, but accounts for only a very small proportion of the variance in PQ scores. The amount of presence reported in a VE is primarily a function of the characteristics of the VE and less a function of individual tendencies. If you wish to measure these individual tendencies, however, the ITQ is recommended.

The PQ is probably the most tested measure of presence that has been developed. It has demonstrated reliability and construct validity and is recommended as the best measure of presence in VE.

VIII Research Issues

The research on the use of VE for dismounted soldier simulation that has been conducted has identified a number of new research issues. The more important ones are described briefly below.

Visual Display Field of View and Performance

Based on our observations, small FOV display devices increase disorientation and reduce the efficiency of spatial learning in VE, and probably are responsible for poor distance perception in VE. Research is needed to confirm these observations, and to determine the most appropriate FOV for dismounted soldier training. Research investigating FOV should use a single display device in which the FOV is manipulated rather than comparing the effects of FOV across display devices.

Separation of Orientation and Distance in Locomotion Simulation

There are methods for locomotion that maintain a constant relationship between real world and VE orientation (Templeman, 1997; Grant and Magee, 1997; Singer, et al., 1998; Lockheed Martin, 1997). The user turns in place in the VE as he or she would in the real world, and therefore receives the same kinesthetic and vestibular cues as in the real world. This aspect of the interface can be used in conjunction with a variety of techniques for moving forward, backward, or sideways. There is no research which assess the contributions of these more realistic cues to distance estimation, route learning, or configuration learning.

Cognitive Demand of VE Interfaces

The cognitive demand imposed by various VE interfaces requires investigation. No research has been conducted which looks at the extent to which the different methods of interaction demand the attention of the user and, therefore, are likely to detract from training effectiveness. Locomotion is the major interface that should be investigated, but certainly not the only one. No research has been conducted which looks at the extent to which the different methods of interaction demand the attention of the user and, therefore, are likely to detract from training effectiveness.

Compensatory Distance Cues

Compensatory auditory cues improved distance estimation performance in one experiment (Witmer and Kline, 1998). It is not known how long the effects of these cues would persist if they were removed, or how well they would transfer to different tasks in the same environment, or to similar environments.

Training Strategies for Spatial Learning

Darken and Siebert (1993, 1996) and Wilson et al. (1997) have shown that VEs can effectively train configuration knowledge, but additional work is needed to determine how VEs can be used to their best advantage in training this type of spatial knowledge. While wider FOV VE displays and more natural VE movement paradigms would alleviate some of the spatial disorientation that occurs in VEs, other alternatives for reducing VE disorientation should be investigated.

VE Technology Improvements and Simulator Sickness

Published recommendations for alleviating simulator sickness have for the most part been directed towards flight simulators. Kennedy et al., (1988) listed several guidelines or rules for reducing simulator sickness which have been implemented at Navy flight training sites. They pointed out that persons most susceptible to simulator sickness are those new to the simulator, and that adaptation of the individual is one of the strongest and most potent fixes for simulator sickness.

For the near-term, we predict that the problem of simulator sickness in VE will parallel the problems of simulator sickness in previous generations of training simulators. Most people will experience a few mild symptoms; some people will experience symptoms severe enough to interfere with training; and for a few people, simulator sickness will be too severe for them to tolerate even brief training sessions in the VE. The threat of simulator sickness in VE training programs will represent an unwanted "overhead" in the form of: additional safety precautions and legal concerns, limits to the frequency and duration of training sessions, restrictions on post-training activities, and a need to provide alternative training for individuals who can not tolerate VEs.

For the long term, a critical question is: Will the problem get worse as the VE experience approaches, but does not equal, real world experience or, will the problem go away as VE technologies improve? McCauley and Sharkey (1992) proposed that simulator sickness is inevitable for a substantial proportion of users of flight and driver simulators. They stated that engineering fixes to simulator sickness are already in the region of diminishing returns. However, they noted that although even excellent engineering may not prevent sickness, poor engineering or calibration will contribute to simulator sickness.

If it turns out that simulator sickness gets worse as the VE experience gets closer to, but does not equal, real-world experience then we should specifically design and conduct research to address ways to prevent or mitigate simulator sickness. In addition, we need to consider if there are serious health or safety issues beyond the discomfort experienced during and immediately after immersion in VEs. That is, are there prolonged or recurring problems associated with VEs? That line of research would require medical expertise in addition to human factors considerations. We will also need a more precise (relative to what we have done before) definition and measurement of system, task, and user characteristics. Finally, greater consideration will be given to pharmacological interventions for treating simulator sickness.

However, simulator sickness may decrease as a function of improvements in VE technologies. For example, future head position tracking systems should have less lag and noise than present systems. In addition, hardware and software innovations may mitigate or eliminate accommodation-vergence conflicts in HMDs. If so, an appropriate course of action is to conduct research for which the emphasis is on training applications of VE and not on simulator sickness per se. For this research we would still measure the occurrence of symptoms and take steps to mitigate simulator sickness during research and training applications.

Presence and Training Effectiveness

Evidence linking the concept of presence to either interface characteristics, on one hand, or training effectiveness, or the other, is weak. In particular, there is no evidence that a trainee who feels highly immersed in his environment learns more or faster than one who does not. However, it is necessary to make a distinction between the level of presence reported by different individuals in response to being immersed in the same environment (as represented by correlating PQ scores and performance) and different amounts of presence reported by the same individual in different environments. The relationship between the presence and task performance in the latter case is yet to be tested.

Minimum Acceptable Update Rate

Aircraft simulators usually update their visual displays at 60 hertz. In VEs, anecdotal evidence and our own experience suggests that "jumps" in the visual scene below the 15-20 hertz range become distracting and annoying. Performance on tracking tasks deteriorates as the update rate decreases (system lag increases). Update rate can usually be increased at the cost of visual detail. Minimum acceptable update rate will probably be based on an interaction between the scene complexity and task-based movement requirements. This trade-off needs to be further investigated in order to be able to generate some "rule of thumb" about edge or flow rates.

Use of Medications to Reduce Simulator Sickness

At least one medicine has been shown to reduce reports of simulator sickness. Quite possibly others do as well. However, the side-effects of these anti-motion sickness medications frequently include drowsiness. They may therefore reduce training effectiveness and transfer if used during the training session. Use of these medications also raises the possibility of a state dependent learning effect (Overton, 1964), i.e., impaired recall of what was originally learned in the medicated state when no longer under the influence of those medications. It is not clear what the best preventative strategy might be for reducing simulator sickness. Is it medication, adaptation without medication, or some combination of the two?

Adaptation and Simulator Sickness

Can users adapt to long-term or repeated use of VEs without experiencing problems adapting to the real world? Some researchers have suggested that extended or repeated effects

can produce delayed flashbacks (Wright, 1995) and sensory-motor disturbances (Stanney and Kennedy, 1997). These require further investigation.

Collective Training in VE

Although our research to date has focused on training individual skills, our ultimate objective is to provide collective (team) and leader training for small units. This will require either training multiple soldiers simultaneously in a shared VE, or training one soldier (the leader) while simulating subordinates and peers, probably with computer-generated forces. Either option requires resolution of a number of additional issues, including: the selection and use of appropriate training features and strategies; determining the extent of system lags (fixed and variable) and their impact on training effectiveness; identifying the requirements for the avatars that are used to represent team members in VE; and determining the impact of reduced non-verbal communication (posture, facial expression) among team members on performance and training effectiveness.

Small Unit Leader Training

In addition to the problems of training teams in VE, there are additional issues which arise from the need to train leader skills that improve performance in combat. Can leaders be trained successfully with simulated team members? Given the current limitations of VE interfaces discussed earlier, (e.g., field of view, lack of inability to aim weapons accurately) can training be provided which is sufficiently similar to combat activities to transfer successfully? Finally, can general team leader skills, as well as the skills which comprise specific unit tasks, be trained?

References

Allen, R C., McDonald, D. P. & Singer, M. J. (1997). Landmark direction and distance estimation in large-scale virtual environments. Proceeding of the Human Factors and Ergonomics Society 41st Annual Meeting,

American Academy of Otolaryngology. (1995). Dizziness and motion sickness [Brochure]. Alexandria, VA: Author.

Anastasi, A. (1968). Psychological testing. New York: McMillan Publishing Company.

Bandura, A. (1971). Social learning theory. Morristown, N.J.: General Learning Corporation.

Bailey, J. H., & Witmer, B. G., (1994). Learning and transfer of spatial knowledge in a virtual environment. Proceedings of the Human Factors & Ergonomics Society 38th Annual Meeting, 1158-1162.

Bliss, J. P., Tidwell, P. D., & Guest, M. A. (1997). The effectiveness of virtual reality for administering spatial navigation training to firefighters. Presence: Teleoperators and Virtual Environments 6(1), 73-86.

Biocca, F. Will simulator sickness slow down the diffusion of virtual environment technology? PRESENCE, 1 (3), 258-264.

Boff, K.R., & Lincoln, J.E. (1988). Engineering data compendium: Human Perception and performance. Wright-Patterson Air Force Base, OH: AAMRL.

Boldovici, J. A. (1992). Toward a Theory of Adaptive Training. (Technical Report 959). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. (ADA251866)

Bowman, D., Koller, D., and Hodges, L. (1997). Travel in immersive virtual environments: an evaluation of viewpoint motion control techniques. Proceedings of the Virtual Reality Annual International Symposium (VRAIS), 45-52.

Cutting, J.E. & Vishton, P.M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein and S.J. Rogers, (Eds.), Handbook of perception and cognition: Volume 5, Perception of space and motion (pp. 69-117). New York: Academic Press.

Darken, R. P. & Sibert, J. L. (1993). A toolset for navigation in virtual environments. Proceedings of ACM User Interface Software and Technology, 157-165.

Darken, R. P. & Sibert, J. L. (1996). Navigating large virtual spaces. International Journal of Human Computer Interaction 8 (1), 49-72.

DiZio, P, & Lackner, J. R., (1992). Spatial orientation, adaptation, and motion sickness in real and virtual environments. Presence, 1(3), 319-328.

Draper, M. H. (1996). Exploring the Influence of a Virtual Body on Spatial Awareness. [On-Line; October, 1997], http://www.hitl.washington.edu/publications/draper/.

Draper, M. H., Wells, M. J., Gawron, V. J., & Furness, T. A. (1996). Exploring the Influence of a Virtual Body on Spatial Awareness, Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting, 1146-1150.

Durlach, N.I., & Mavor, A.S. (Eds.). (1995). Virtual reality: scientific and technological challenges. Washington DC: National Academy Press.

Dyer, R., Matthews, J. J., Stulac, J. F., Wright, C. E., and Yudowitch, K. (1976). Questionnaire construction manual, annex literature survey and bibliography. Palo Alto, CA: Operations Research Associates.

England (1995). Sensory-motor systems in virtual manipulation. in eds. Carr, K & England, R. Simulated and Virtual Realities.

Garris-Reif, R., & Franz, T.M. (1995). Simulator sickness and human task performance in conventional and virtual environments. In A.C. Bittner & P.C. Champney (Eds.), Advances in Industrial Ergonomics and Safety VII. Seattle, WA: Taylor & Francis.

Gibson, E. J. (1969). Principles of perceptual learning and development. Englewood Cliffs, N.J.: Prentice-Hall.

Goldberg, S.L., & Knerr, B.W. (1997). Collective training in virtual environments: Exploring performance requirements for dismounted soldier simulation. In R.J. Seidel & P.R. Chatelier (Eds.), Virtual reality, training's future? New York: Plenum Press.

Goldin, S. E., & Thorndyke, P. W. (1982). Simulating navigation for spatial knowledge acquisition. Human Factors, 24(4), 457-471.

Gorman, P.F. (1990). Supertroop via I-Port: Distributed simulation technology for combat development and training development (IDA Paper P-2374). Alexandria, VA: Institute for Defense Analysis.

Grant, S.C., & Magee, L.E. (1997, December). Navigation in a virtual environment using a walking interface. Paper presented at the NATO RSG-18 Workshop on Capability of Virtual reality to Meet Military Requirements. Orlando, FL.

Hays, R. T., & Singer, M. J. (1989). Simulation Fidelity in Training System Design: Bridging the Gap between Reality and Training. New York: Springer-Verlag.

Held, R. & Durlach, N. (1992). Telepresence. Presence, 1(1), 109-112.

Hettinger, L. J. & Riccio, G. E., (1992). Visually induced motion sickness in virtual environments. Presence, 1 (3), 306-310.

Jacobs, R. S., Crooks, W. H., Crooks, J. R., Colburn, E., Fraser, II, R. E., Gorman, P. F., Madden, J. L., Furness, III, T. A., & Tice, S. E. (October, 1994). Behavioral Requirements for Training and Rehearsal in Virtual Environments. (TR-1011), U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA. (ADA286311)

Johnson, D. M. & Wightman, D. C. (1995). Using virtual environments for terrain familiarization. (Research Report 1686). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA304416)

Kaufman, J. E., & Christiansen J. F., (Eds.) (1984). Light and Vision. Illuminating Engineering Society Lighting Handbook: 1984 Reference Volume, New York, Illuminating Engineering Society.

Kennedy, R.S., Berbaum, K. S., Lilienthal, M.G., Dunlap, D.W., Mulligan, B.F., & Funaro, J.F., (1987, March). Guidelines for alleviation of simulator sickness symptomatology (NAVTRASYSCEN TR-87007). Orlando, FL: Naval Training Systems Center. (AD A182-554)

Kennedy, R. S., & Fowlkes, J.E. (1992) Simulator sickness is polygenic and polysymptomatic: Implications for research. International Journal of Aviation Psychology, 2(1), 23-38.

Kennedy, R. S., Fowlkes, J.E., Berbaum, K. S., & Lilienthal, M.G., (1992). Use of a motion sickness history questionnaire for prediction of simulator sickness. Aviation, Space, and Environmental Medicine, 63(7), 588-593.

Kennedy, R.S., Hettinger, L.J., & Lilienthal, M.G (1988). Simulator Sickness. In G.H. Crampton (Ed.), Motion and Space Sickness. 317-341. Boca Raton, FL: CRC Press.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M.G., (1993). A simulator sickness questionnaire (SSQ): A new method for quantifying simulator sickness. International Journal of Aviation Psychology. 3(3), 203-220.

Kline, P. B. & Witmer, B. G. (1996). Distance perception in virtual environments: Effects of field of view and surface texture at near distances. Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting, 1112-1116.

Knerr, B. W., Goldberg, S. L., Lampton, D. R., Witmer, B. G., Bliss, J. P., Moshell, J. M., & Blau, B. S. (1994). Research in the use of virtual environment technology to train dismounted soldiers. Journal of Interactive Instruction Development, 6(4), 9-20.

Kolasinski, E.M. (1995). Simulator sickness in virtual environments (Technical Report 1027). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA295861)

Kozak, J. J., Hancock, P. A., Arthur, E. J., & Chrysler, S. T. (1993). Transfer of training from virtual reality. Ergonomics, 36(7), 777-784.

Lampton, D. L., Gildea, J. P., McDonald, D. P. & Kolasinski, E. U. (1996). Effects of display type on performance in virtual environments. (Technical Report 1049). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA322046)

Lampton, D. L., Knerr, B. W., Goldberg, S. L., Bliss, J. P., Moshell, J. M., & Blau, B. (1994). The virtual environment performance assessment battery (VEPAB): Development and evaluation. PRESENCE 3(2), 145-157.

Lampton, D. R., Knerr, B. W., Goldberg, S.L., Bliss, J. P., Moshell, J. M., & Blau, B.S. (1995). The virtual environment performance assessment battery (VEPAB): Development and evaluation (Technical Report 1029). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA292277)

Lampton, D.L., Knerr, B.W., McDonald, D., Rodriguez, M., and Parsons, J. (1998). Instructional Strategies for Team Training in Virtual Environments. Submitted for publication in the proceedings of the Image Society 1998 biannual meeting.

Lampton, D.R., Kraemer, R.E., Kolasinski, E., M., & Knerr, B.W. (1995). An investigation of simulator sickness in a tank driver trainer (Research Report 1684). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences. (ADA304277)

Lampton, D. L., McDonald, D. P., & Singer, M. (1995). Distance estimation in virtual environments. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting, 1268-1272.

Levison, W. H., & Pew, R. W. (1993). Use of virtual environment training technology for individual combat simulation. (Technical Report 971). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA263546)

Lockheed Martin Information Systems (1997). Dismounted warrior network front-end analysis experiments final report. Orlando, FL: Author.

Magee, L.M. (1997). Virtual Reality Simulator (VRS) for Training Ship Handling Skills. In Robert J. Seidel and Paul R. Chatelier (Eds) Virtual Reality, Training's Futur:e: Perspectives on Virtual Reality and Related Emerging Technologies. New York: Plenum Press.

McCauley, M.E., & Sharkey, T.J. (1992). Cybersickness: Perception of self-motion in virtual environments. PRESENCE, 1(3), 311-318.

Moffit, K. W. (1997). Designing HMDs for viewing comfort In J.E. Melzer, & K.W. Moffit, (Eds.), Head-Mounted Displays: Designing for the User (pp.117-145). New York: McGraw Hill.

Naylor, S. D. (1996, June 10). 'Next' up: Looking beyond Force XII. Army Times, p. 18.

National Research Council (1997). Tactical display for soldiers: Human factors considerations. Washington, D.C.: National Academy Press.

Overton, D.A. (1964). State-dependent or 'dissociate' learning produced with pentobarbitol. Journal of Comparative and Physiological Psychology, 57(1), 3-12.

Pleban, R. J., Dyer, J. L., Salter, M. S., & Brown, J. B. (1997). Functional capabilities of four virtual individual combatant (VIC) simulator technologies: An independent assessment. (Technical report 1078). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences. (ADA343575)

Regan, C. (1995) An investigation into nausea and other side-effects of head-coupled immersive virtual reality. Virtual Reality, 1(1), 17-32.

Regan, E.C., & Price, K.R. (1993). Some side effects of immersion in virtual reality: An investigation into the relationship between interpupillary distance and ocular related problems (APRE Report 93R023). Farnsborough, Hants, UK: Army Personnel Research Establishment.

Regian, J. W., Shebilske, W. L., & Monk, J. M. (1992). Virtual reality: An instructional medium for visual-spatial tasks. Journal of Communication, 42(4), 136-149.

Rinalducci, E. J., Cinq-Mars, S. C., Mapes, D., & Higgins, K. E. (1996). Determining the field of view in HMDs: A psychophysical method. PRESENCE, 5(3), 353-356.

Seligman, M. E. P. (1970). On the generality of the law of learning. Psychological Review, 77, 406-418.

Semple, C. A., Cotton, J. C., & Sullivan, D. J. (1981). Aircrew Training Devices: Instructional Support Features, (AFHRL-TR-80-58). Wright-Patterson AFB, Ohio: U. S. Air Force Human Resources Laboratory. (ADA 096234)

Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. Presence, 1(1), 120-125.

Siegal, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In Reese, H. (Ed.), Advances in Child Development and Behavior, Vol. 10, New York: Academic Press, 10-55.

Singer, M. J., Allen, R. C., McDonald, D. P., & Fober, G. (in preparation). Acquisition and transfer of spatial knowledge from maps versus virtual environments. (Draft Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Singer, M. J., Allen, R. C., McDonald, D. P., & Gildea, J. P. (1997). Terrain appreciation in virtual environments: Spatial knowledge and acquisition. (Technical Report 1056). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA325520)

Singer, M. J., Ehrlich, E. A., & Allen, R. C. (1998). Effect of a Body Model on Performance in a Virtual Environment Search Task. (Technical Report 1087). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. (ADA352026)

Singer, M. J., Ehrlich, J., & Cinq-Mars, S. & Papin, Jean-Paul (1995). Task performance in virtual environments: Stereoscopic versus monoscopic displays and head-coupling. (Technical Report 1034). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. (ADA306720)

Slater, M., Linakis, V., Usoh, M. & Kooper, R. (1996, July). Immersion, presence, and performance in virtual environments: An experiment using tri-dimensional chess. Paper presented at the 1996 Virtual Reality and Software and Technology Conference, Hong Kong.

Soldier Systems Command. (1998). Products and programs [on-line]. Available www-sscom.army.mil/prod/prog/index/html

Sollenberger, R. L., & Milgram, P. (1993). Effects of stereoscopic and rotational displays in a three-dimensional path-tracing task. Human Factors, 35(3), 483-499.

Stanney, K., & Kennedy, R.S. (1997). Development and testing of a measure of the kinesthetic position sense used to assess the aftereffects from virtual environment exposure. Proceedings of the VR Annual International Symposium, 87-94. Albuquerque, NM.

Stanney, K., Salvendy, G., Deisinger, J., Ellis, S., Ellison, J., Fogelman, G., Gallimore, J. Hettinger, L., Kennedy, R., Lackner, J., Lawson, B., Maida, J, Mead, A., Mon-Williams, M., Newman, D., Plantinida, T., Reeves, L., Riedel, O., Singer, M., Stoffregan, T. Wann, J., Welch, R., Wilson, J., and Witmer, B. (in press). Aftereffects and sense of presence in virtual environments: Development of an R&D agenda. International Journal of Human Computer Interaction.

Surdick, R.T., Davis, E., King, R.A., and Hodges, L.F. (1997). The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances. PRESENCE: Teleoperators and Virtual Environments 6(5), 513-531.

Templeman, J. (1997, December). Performance based design of a new virtual locomotion control. Paper presented at the NATO RSG-18 Workshop on Capability of Virtual Reality to Meet Military Requirements. Orlando, FL.

Thomley, K.E., Kennedy, R.S., & Bittner, A.C. (1986). Development of postural equilibrium tests for examining environmental effects. Perceptual and Motor Skills, 63, 555-564.

Thorndyke, P. W., & Hayes-Roth, B. (1980). Differences in spatial knowledge acquired from maps and navigation. (Report N-1595-ONR). Santa Monica, CA: The Rand Corporation.

Triesman, M. (1977). Motion sickness: An evolutionary hypothesis. Science, 197, 493-495.

Tulving, E. (1983). Elements of Episodic Memory. Oxford: Oxford University Press.

Underwood, B. J. & Schulz, R. W. (1960). Meaningfulness and verbal learning. Philadelphia, Pa.: Lippincott.

Wilson, P. N., Foreman, N., & Tlauka, M. (1997). Transfer of spatial information from a virtual to a real environment. Human Factors 39(4), 526-531.

Witmer, B. G., Bailey, J. H., & Knerr, B. W. (1995). Training dismounted soldiers in virtual environments: Route learning and transfer. (Technical Report 1022). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Witmer, B. G., Bailey, J. H., Knerr, B. W., & Parsons, K. C. (1996). Virtual spaces and real world places: Transfer of route knowledge. International Journal of Human Computer Studies 45, 413-428.

Witmer, B. G. & Kline, P. B. (1998). Judging perceived and traversed distance in virtual environments. PRESENCE: Teleoperators and Virtual Environments 7(2), 144-167.

Witmer, B. G. & Sadowski, W. J. (in press). Non-visually guided locomotion to previously viewed target in real and virtual environments. Human Factors.

Witmer, B. G., & Singer, M. J. (1994). Measuring presence in virtual environments. (Technical Report 1014). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. (ADA 286183)

Witmer, B.G. & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. Presence: Teleoperators and Virtual Environments, 7(3), 225-240.

Wood, J. M., & Troutbeck, R. (1992). Effect of restriction of the binocular visual field on driving performance, Ophthalmology Physiological Optics, 12, 291-298.

Wright, R. H. (1995). Virtual reality psychophysics: Forward and lateral distance, height and speed perceptions with a wide angle helmet display. (Technical Report 1025). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences. (ADA294027)

Wright, R. H. (1995). Helicopter simulator sickness: A state-of-the-art review of its incidence, causes, and treatment (Research Report 1680). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences. (ADA297285)

APPENDIX A

ACRONYMS AND ABBREVIATIONS

AGP ....Advanced Graphics Port
API ....Application Programming Interface
ARI .................................U.S. Army Research Institute for the Behavioral and Social Sciences
BOOM (or BOOM2C).....Binocular Omni-Orientation Monitor
CPU ....Central Processing Unit
DI ....Dismounted Infantry
DIS ....Distributed Interactive Simulation
DWN ....Dismounted Warrior Network
FOV ....Field of View
GFOV ....Geometric Field of View
HMD ....Helmet-Mounted Display or Head-Mounted Display
ICSS ....Individual Combatant Simulation System
I-Port ....Individual Portal into VE
IST ....University of Central Florida Institute for Simulation and Training
ITQ ....Immersive Tendencies Questionnaire
M1TDT ....M1 Tank Driver Trainer
MARS VRS ....Maritime Surface/Subsurface Virtual Reality Simulator
MOUT ....Military Operations in Urban Terrain
NVGL ....Non-visually Guided Locomotion
PDU ....Protocol Data Unit
PQ ....Presence Questionnaire
RW ....Real World
SAF ....Semi-Automated Forces
SGI ....Silicon Graphics, Inc.
SSQ ....Simulator Sickness Questionnaire
STRICOM .................................U.S. Army Simulation Training and Instrumentation Command
TRADOC .................................U.S. Army Training and Doctrine Command
VE ....Virtual Environment
VEPAB ....Virtual Environment Performance Assessment Battery
VIC ....Virtual Individual Combatant
VIMS ....Visually Induced Motion Sickness